

IBM System z10 Enterprise Class (z10 EC) Reference Guide



Table of Contents

<i>IBM System z10 EnterpriseClass (z10 EC) Overview</i>	<i>page 3</i>
<i>z/Architecture</i>	<i>page 6</i>
<i>z10 EC</i>	<i>page 11</i>
<i>z10 EC Design and Technology</i>	<i>page 14</i>
<i>z10 EC Model</i>	<i>page 15</i>
<i>z10 EC Performance</i>	<i>page 17</i>
<i>z10 EC I/O Subsystem</i>	<i>page 18</i>
<i>z10 EC Channels and I/O Connectivity</i>	<i>page 19</i>
<i>HiperSockets</i>	<i>page 32</i>
<i>Security</i>	<i>page 34</i>
<i>Cryptography</i>	<i>page 34</i>
<i>On Demand Capabilities</i>	<i>page 39</i>
<i>Reliability, Availability, and Serviceability (RAS)</i>	<i>page 43</i>
<i>Availability Functions</i>	<i>page 44</i>
<i>Environmental Enhancements</i>	<i>page 47</i>
<i>Parallel Sysplex Cluster Technology</i>	<i>page 48</i>
<i>HMC System Support</i>	<i>page 57</i>
<i>Implementation Services for Parallel Sysplex</i>	<i>page 59</i>
<i>Fiber Quick Connect for FICON LX Environments</i>	<i>page 60</i>
<i>z10 EC Physical Characteristics</i>	<i>page 60</i>
<i>z10 EC Configuration Detail</i>	<i>page 61</i>
<i>Coupling Facility – CF Level of Support</i>	<i>page 64</i>
<i>Statement of Direction</i>	<i>page 65</i>
<i>Publications</i>	<i>page 66</i>

IBM System z10 Enterprise Class (z10 EC) Overview

The IBM System z10™ Enterprise Class (z10™ EC) server is designed to meet the challenges of today's business world and to be the cornerstone of an evolutionary new model for efficient IT delivery called the Dynamic Infrastructure®. This model helps reset the economics of IT and can dramatically improve operational efficiency, security, and responsiveness – to help keep a business competitive.

The z10 EC™, with its advanced combination of reliability, availability, serviceability, security, scalability, and virtualization, delivers the technology that can help define this framework for the future. The z10 EC delivers improvements to performance, capacity, and memory which can help enterprises grow their existing business while providing a cost-effective infrastructure for large-scale consolidation.

The October 2008 announcements extend the z10 EC leadership with improved access to data and the network; tighter security with longer Personal Account Numbers for stronger protection of data; enhancements for improved performance when connecting to the network; increased flexibility in defining your options to handle backup requirements; and enhanced time accuracy to an external time source.

Any successful business needs to be able to deliver timely, integrated information to business leaders, support personnel, and customers on a 24x7 basis. This means that access to data needs to be fast, secure, and dependable. Enhancements made to z/Architecture® and the FICON® interface architecture with the High Performance FICON for System z (zHPF) are optimized for online transaction processing (OLTP) workloads. The FICON Express4 and FICON Express2 features support the native FICON protocol and the zHPF protocol.

The System z10 was introduced with a new connectivity option for LANs – Open Systems Adapter-Express3 (OSA-Express3). The OSA-Express3 features provide improved performance by reducing latency at the TCP/IP application. Direct access to the memory allows packets to flow directly from the memory to the LAN without firmware intervention in the adapter.

An IT system needs to be available and protected every day. The z10 EC offers availability enhancements which include faster service time for CF Duplexing, updates to Server Time Protocol (STP) for enhanced time accuracy to an External Time Source, and support for heterogeneous platforms in an enterprise to track to the same time source. Security enhancements to the Crypto Express2 feature deliver support for 13-, 14-, 15-, 16-, 17-, 18-, and 19-digit Personal Account Numbers for stronger protection of data.

The z10 EC has a new architectural approach for temporary offerings that have the potential to change the thinking about on demand capacity. The z10 EC can have one or more flexible configuration definitions that can be available to solve multiple temporary situations and multiple capacity configurations that can be active at once. This means that On/Off Capacity on Demand (CoD) can be active and up to seven other offerings can be active simultaneously. Tokens are available that can be purchased for On/Off CoD either before or after execution.

Updates to the z10 EC are designed to help improve IT today, outline a compelling case for the future running on System z, and lock in the z10 EC as the cornerstone in your Dynamic Infrastructure by delivering superior business and IT services with agility and speed.

Just-in-time deployment of IT resources

Infrastructures must be more flexible to changing capacity requirements and provide users with just-in-time deployment of resources. Having the 16 GB dedicated HSA on the z10 EC means that some preplanning configuration changes and associated outages may be avoided. IBM Capacity Upgrade on Demand (CUoD) provides a permanent increase in processing capacity that can be initiated by the customer.

IBM On/Off Capacity on Demand (On/Off CoD) provides temporary capacity needed for short-term spikes in capacity or for testing new applications. Capacity Backup Upgrade (CBU) can help provide reserved emergency backup capacity for all processor configurations.

An additional temporary capacity offering on the z10 EC is Capacity for Planned Events (CPE), a variation on CBU. If unallocated capacity is available in a server, it will allow the maximum capacity available to be used for planned events such as planned maintenance in a data center.

By having flexible and dynamic configuration definitions, when capacity is needed, activation of any portion of an offering can be done (for example activation of just two CBUs out of a definition that has four CBUs is acceptable). And if the definition doesn't have enough resources defined, an order can easily be processed to increase the capacity (so if four CBUs aren't enough it can be redefined to be six CBUs) as long as enough server infrastructure is available to meet maximum needs.

All activations can be done without having to interact with IBM—when it is determined that capacity is required, no passwords or phone connections are necessary. As long as the total z10 EC can support the maximums that are defined, then they can be made available.

With the z10 EC, it is now possible to add permanent capacity while a temporary capacity is currently activated, without having to return first to the original configuration.

The activation of On/Off CoD on z10 EC can be simplified or automated by using z/OS Capacity Provisioning (available with z/OS® 1.9 and above). This capability enables the monitoring of multiple systems based on Capacity Provisioning and Workload Manager (WLM) definitions. When the defined conditions are met, z/OS can suggest capacity changes for manual activation from a z/OS console, or the system can add or remove temporary capacity automatically and without operator intervention.

Specialty engines offer an attractive alternative

The z10 EC continues the long history of providing integrated technologies to optimize a variety of workloads. The use of specialty engines can help users expand the use of the mainframe for new workloads, while helping to lower the cost of ownership. The IBM System z® specialty engines can run independently or complement each other. For example, the zAAP and zIIP processors enable you to purchase additional processing capacity exclusively for specific workloads, without affecting the MSU rating of the IBM System z model designation. This means that adding a specialty engine will not cause increased charges for IBM System z software running on general purpose processors in the server.

In order of introduction:

The Internal Coupling Facility (ICF) processor was introduced to help cut the cost of Coupling Facility functions by reducing the need for an external Coupling Facility. IBM System z Parallel Sysplex® technology allows for greater scalability and availability by coupling mainframes together. Using Parallel Sysplex clustering, System z servers are designed for up to 99.999% availability.

The Integrated Facility for Linux (IFL) processor offers support for Linux® and brings a wealth of available applications that can be run in a real or virtual environment on the z10 EC. An example is the z/VSE™ strategy which supports integration between the IFL, z/VSE and Linux on System z to help customers integrate timely production of z/VSE data into new Linux applications, such as data warehouse environments built upon a DB2® data server. To consolidate distributed servers onto System z, the IFL with Linux and the System z virtualization technologies fulfill the qualifications for business-critical workloads as well as for infrastructure workloads. For customers interested to use a z10 EC only for Linux workload, the z10 EC can be configured as a server with IFLs only.

Available on System z since 2004, the System z10 Application Assist Processor (zAAP) is designed to help enable strategic integration of new application technologies such as Java™ technology-based Web applications and XML-based data interchange services with core business database environments. This helps provide a more cost-effective, specialized z/OS application Java execution environment. Workloads eligible for the zAAP (with z/OS V1.8) include all Java processed via the IBM Solution Developers Kit (SDK) and XML processed locally via z/OS XML System Services.

The System z10 Integrated Information Processor (zIIP) is designed to support select data and transaction processing and network workloads and thereby make the consolidation of these workloads on to the System z platform more cost effective. Workloads eligible for the zIIP (with z/OS V1.7 or later) include remote connectivity to DB2 to help support these workloads: Business Intelligence (BI), Enterprise Relationship Management (ERP), Customer Relationship Management (CRM) and Extensible Markup Language (XML) applications. In addition to supporting remote connectivity to DB2 (via DRDA® over TCP/IP) the zIIP also supports DB2 long running parallel queries—a workload integral to Business Intelligence and Data Warehousing solutions. The zIIP (with z/OS V1.8) also supports IPSec processing, making the zIIP an IPSec encryption engine helpful in creating highly secure connections in an enterprise. In addition, zIIP (with z/OS V1.10) supports select z/OS Global Mirror (formerly called Extended Remote Copy, XRC) disk copy service functions. z/OS V1.10 also introduces zIIP-Assisted HiperSockets™ for large messages (available on System z10 servers only).

The new capability provided with z/VM®-Mode partitions increases flexibility and simplifies systems management by allowing z/VM 5.4 to manage guests to operate Linux on System z on IFLs, to operate z/VSE and z/OS on CPs, to offload z/OS system software overhead, such as DB2 workloads on zIIPs, and to offer an economical Java execution environment under z/OS on zAAPs, all in the same z/VM LPAR.

Numerical computing on the chip

Integrated on the z10 EC processor unit is a Hardware Decimal Floating Point unit to accelerate decimal floating point transactions. This function is designed to markedly improve performance for decimal floating point operations which offer increased precision compared to binary floating

z/Architecture

point operations. This is expected to be particularly useful for the calculations involved in many financial transactions.

Decimal calculations are often used in financial applications and those done using other floating point facilities have typically been performed by software through the use of libraries. With a hardware decimal floating point unit, some of these calculations may be done directly and accelerated.

Liberating your assets with System z

Enterprises have millions of dollars worth of mainframe assets and core business applications that support the heart of the business. The convergence of service oriented architecture (SOA) and mainframe technologies can help liberate these core business assets by making it easier to enrich, modernize, extend and reuse them well beyond their original scope of design. The z10 EC, along with the inherent strengths and capabilities of a z/OS environment, provides an excellent platform for being an enterprise hub. Innovative System z software solutions from WebSphere®, CICS®, Rational® and Lotus® strengthen the flexibility of doing SOA.

Evolving for your business

The z10 EC is the next step in the evolution of the System z mainframe, fulfilling our promise to deliver technology improvements in areas that the mainframe excels in—energy efficiency, scalability, virtualization, security and availability. The redesigned processor chip helps the z10 EC make high performance compute-intensive processing a reality. Flexibility and control over capacity gives IT the upper edge over planned or unforeseen demands. And new technologies can benefit from the inherent strengths of the mainframe. This evolving technology delivers a compelling case for the future to run on System z.

The z10 EC continues the line of upward compatible mainframe processors and retains application compatibility since 1964. The z10 EC supports all z/Architecture-compliant Operating Systems. The heart of the processor unit is the Enterprise Quad Core z10 Processor Unit chip which is specifically designed and optimized for mainframe systems. New features enhance enterprise data serving performance as well as CPU-intensive workloads.

The z10 EC, like its predecessors, supports 24-, 31-, and 64-bit addressing, as well as multiple arithmetic formats. High-performance logical partitioning via Processor Resource/Systems Manager™ (PR/SM™) is achieved by industry-leading virtualization support provided by z/VM.

z10 EC Architecture

Rich CISC Instruction Set Architecture (ISA)

- *894 instructions (668 implemented entirely in hardware)*
- *Multiple address spaces robust inter-process security*
- *Multiple arithmetic formats*

Architectural extensions for z10 EC

- *50+ instructions added to z10 EC to improve compiled code efficiency*
- *Enablement for software/hardware cache optimization*
- *Support for 1 MB page frames*
- *Full hardware support for Hardware Decimal Floating-point Unit (HDFU)*

z/Architecture operating system support

Delivering the technologies required to address today's IT challenges also takes much more than just a server; it requires all of the system elements to be working together.

IBM system z10 operating systems and servers are designed with a collaborative approach to exploit each other's strengths.

The z10 EC is also able to exploit numerous operating systems concurrently on a single server, these include z/OS, z/VM, z/VSE, z/TPF, TPF and Linux for System z. These operating systems are designed to support existing application investments without anticipated change and help you realize the benefits of the z10 EC. System z10 – the new business equation.

z/OS

August 5, 2008, IBM announced z/OS V1.10. This release of the z/OS operating system builds on leadership capabilities, enhances time-tested technologies, and leverages deep synergies with the IBM System z10 and IBM System Storage™ family of products. z/OS V1.10 supports new capabilities designed to provide:

- *Storage scalability. Extended Address Volumes (EAVs) enable you to define volumes as large as 223 GB to relieve storage constraints and help you simplify storage management by providing the ability to manage fewer, large volumes as opposed to many small volumes.*
- *Application and data serving scalability. Up to 64 engines, up to 1.5 TB per server with up to 1.0 TB of real memory per LPAR, and support for large (1 MB) pages on the System z10 can help provide scale and performance for your critical workloads.*
- *Intelligent and optimized dispatching of workloads. Hip-erDispatch can help provide increased scalability and performance of higher n-way z10 EC systems by improving the way workload is dispatched within the server.*
- *Low-cost, high-availability disk solution. The Basic HyperSwap™ capability (enabled by TotalStorage® Productivity Center for Replication Basic Edition for System z) provides a low-cost, single-site, high-availability disk solution which allows the configuration of disk replication services using an intuitive browser-based graphical user interface (GUI) served from z/OS.*

- *Improved total cost of ownership. zIIP-Assisted HiperSockets for Large Messages, IBM Scalable Architecture for Financial Reporting™ enabled for zIIP (a service offering of IBM Global Business Services), zIIP-Assisted z/OS Global Mirror (XRC), and additional z/OS XML System Services exploitation of zIIP and zAAP help make these workloads more attractive on System z.*
- *Improved management of temporary processor capacity. A Capacity Provisioning Manager, which is available on z/OS V1.10, and available on z/OS V1.9 with PTFs, can monitor z/OS systems on z10 EC servers. Activation and deactivation of temporary capacity can be suggested or performed automatically based on user-defined schedules and workload criteria. RMF™ or equivalent function is required to use the Capacity Provisioning Manager.*
- *Improved network security. z/OS Communications Server introduces new defensive filtering capability. Defensive filters are evaluated ahead of configured IP filters, and can be created dynamically, which can provide added protection and minimal disruption of services in the event of an attack.*
- *z/OS V1.10 also supports RSA key, ISO Format-3 PIN block, 13-Digit through 19-Digit PANdata, secure key AES, and SHA algorithms.*
- *Improved productivity. z/OS V1.10 provides improvements in or new capabilities for: simplifying diagnosis and problem determination; expanded Health Check Services; network and security management; automatic dump and re-IPL capability; as well as overall z/OS, I/O configuration, sysplex, and storage operations*

With z/OS 1.9, IBM delivers functionality that continues to solidify System z leadership as the premier data server. z/OS 1.9 offers enhancements in the areas of security, networking, scalability, availability, application development, integration, and improved economics with more exploitation for specialty engines. A foundational element of the platform — the z/OS tight interaction with the System z hardware and its high level of system integrity.

With z/OS 1.9, IBM introduces:

- *A revised and expanded Statement of z/OS System Integrity*
- *Large Page Support (1 MB)*
- *Capacity Provisioning*
- *Support for up to 64 engines in a single image (on z10 EC model only)*
- *Simplified and centralized policy-based networking*
- *Expanded IBM Health Checker*
- *Simplified RACF® Administration*
- *Hardware Decimal Floating Point*
- *Parallel Sysplex support for Infiniband® Coupling Links*
- *NTP Support for STP*
- *HiperSockets Multiple Write Facility*
- *OSA-Express3 support*
- *Advancements in ease of use for both new and existing IT professionals coming to z/OS*
- *Support for zIIP-Assisted IPSec, System Data Mover (SDM) offload to zIIP, and support for eligible portions of DB2 9 XML parsing workloads to be offloaded to zAAP processors*
- *Expanded options for AT-TLS and System SSL network security*
- *Improved creation and management of digital certificates with RACF, SAF, and z/OS PKI Services*
- *Additional centralized ICSF encryption key management functions for applications*
- *Improved availability with Parallel Sysplex and Coupling Facility improvement*
- *Enhanced application development and integration with new System REXX™ facility, Metal C facility, and z/OS UNIX® System Services commands*
- *Enhanced Workload Manager in managing discretionary work and zIIP and zAAP workloads*

Commitment to system integrity

First issued in 1973, IBM's MVS™ System Integrity Statement and subsequent statements for OS/390® and z/OS stand as a symbol of IBM's confidence and commitment to the z/OS operating system. Today, IBM reaffirms its commitment to z/OS system integrity.

IBM's commitment includes designs and development practices intended to prevent unauthorized application programs, subsystems, and users from bypassing z/OS security—that is, to prevent them from gaining access, circumventing, disabling, altering, or obtaining control of key z/OS system processes and resources unless allowed by the installation. Specifically, z/OS “System Integrity” is defined as the inability of any program not authorized by a mechanism under the installation's control to circumvent or disable store or fetch protection, access a resource protected by the z/OS Security Server (RACF), or obtain control in an authorized state; that is, in supervisor state, with a protection key less than eight (8), or Authorized Program Facility (APF) authorized. In the event that an IBM System Integrity problem is reported, IBM will always take action to resolve it.

IBM's long-term commitment to System Integrity is unique in the industry, and forms the basis of the z/OS industry leadership in system security. z/OS is designed to help you protect your system, data, transactions, and applications from accidental or malicious modification. This is one of the many reasons System z remains the industry's premier data server for mission-critical workloads.

z/VM

z/VM V5.4 is designed to extend its System z virtualization technology leadership by exploiting more capabilities of System z servers including:

- *Greater flexibility, with support for the new z/VM-mode logical partitions, allowing all System z processor-types (CPs, IFLs, zIIPs, zAAPs, and ICFs) to be defined in the same z/VM LPAR for use by various guest operating systems*
- *Capability to install Linux on System z as well as z/VM from the HMC on a System z10 that eliminates the need for any external network setup or a physical connection between an LPAR and the HMC*
- *Enhanced physical connectivity by exploiting all OSA-Express3 ports, helping service the network and reducing the number of required resources.*
- *Dynamic memory upgrade support that allows real memory to be added to a running z/VM system. With z/VM V5.4, memory can be added non-disruptively to individual guests that support the dynamic memory reconfiguration architecture. Systems can now be configured to reduce the need to re-IPL z/VM. Processors, channels, OSA adapters, and now memory can be dynamically added to both the z/VM system itself and to individual guests.*

The operation and management of virtual machines has been enhanced with new systems management APIs, improvements to the algorithm for distributing a guest's CPU share among virtual processors, and usability enhancements for managing a virtual network.

Security capabilities of z/VM V5.4 provide an upgraded LDAP server at the functional level of the z/OS V1.10 IBM Tivoli® Directory Server for z/OS and enhancements to the RACF Security Server to create LDAP change log entries in response to updates to RACF group and user profiles, including user passwords and password phrases. The z/VM

SSL server now operates in a CMS environment, instead of requiring a Linux distribution, thus allowing encryption services to be deployed more quickly and helping to simplify installation, service, and release-to-release migration.

The z/VM hypervisor is designed to help clients extend the business value of mainframe technology across the enterprise by integrating applications and data while providing exceptional levels of availability, security, and operational ease. z/VM virtualization technology is designed to provide the capability for clients to run hundreds to thousands of Linux servers in a single mainframe, together with other System z operating systems such as z/OS, or as a large-scale Linux-only enterprise-server solution. z/VM V5.4 can also help to improve productivity by hosting non-Linux workloads such as z/OS, z/VSE, and z/TPF.

August 5, 2008, IBM announced z/VM 5.4. Enhancements in z/VM 5.4 include:

- *Increased flexibility with support for new z/VM-mode logical partitions*
- *Dynamic addition of memory to an active z/VM LPAR by exploiting System z dynamic storage-reconfiguration capabilities*
- *Enhanced physical connectivity by exploiting all OSA-Express3 ports*
- *Capability to install Linux on System z from the HMC without requiring an external network connection*
- *Enhancements for scalability and constraint relief*
- *Operation of the SSL server in a CMS environment*
- *Systems management enhancements for Linux and other virtual images*

For the most current information on z/VM, refer to the z/VM Web site at <http://www.vm.ibm.com>.

z/VSE

z/VSE 4.1, the latest advance in the ongoing evolution of VSE, is designed to help address needs of VSE clients with growing core VSE workloads and/or those who wish to exploit Linux on System z for new, Web-based business solutions and infrastructure simplification.

z/VSE 4.1 is designed to support:

- *z/Architecture mode only*
- *64-bit real addressing and up to 8 GB of processor storage*
- *System z encryption technology including CPACF, configurable Crypto Express2, and TS1120 encrypting tape*
- *Midrange Workload License Charge (MWLC) pricing, including full-capacity and sub-capacity options.*

IBM has previewed z/VSE 4.2. When available, z/VSE 4.2 is designed help address the needs of VSE clients with growing core VSE workloads. z/VSE V4.2 is designed to support:

- *More than 255 VSE tasks to help clients grow their CICS workloads and to ease migration from CS/VSE to CICS Transaction Server for VSE/ESA™*
- *Up to 32 GB of processor storage*
- *Sub-Capacity Reporting Tool running “natively”*
- *Encryption Facility for z/VSE as an optional priced feature*
- *IBM System Storage TS3400 Tape Library (via the TS1120 Controller)*
- *IBM System Storage TS7740 Virtualization Engine Release 1.3*

z/VSE V4.2 plans to continue the focus on hybrid solutions exploiting z/VSE and Linux on System z, service-oriented architecture (SOA), and security. It is the preferred replacement for z/VSE V4.1, z/VSE V3, or VSE/ESA. It is designed to protect and leverage existing VSE information assets.

z/TPF

z/TPF is a 64-bit operating system that allows you to move legacy applications into an open development environment, leveraging large scale memory spaces for increased speed, diagnostics and functionality. The open development environment allows access to commodity skills and enhanced access to open code libraries, both of which can be used to lower development costs. Large memory spaces can be used to increase both system and application efficiency as I/Os or memory management can be eliminated.

z/TPF is designed to support:

- *64-bit mode*
- *Linux development environment (GCC and HLASM for Linux)*
- *32 processors/cluster*
- *Up to 84* engines/processor*
- *40,000 modules*
- *Workload License Charge*

Linux on System z

The System z10 EC supports the following Linux on System z distributions (most recent service levels):

- *Novell SUSE SLES 9*
- *Novell SUSE SLES 10*
- *Red Hat RHEL 4*
- *Red Hat RHEL 5*

z10 EC

Operating System	ESA/390 (31-bit)	z/Architecture (64-bit)
z/OS V1R8, 9 and 10	No	Yes
z/OS V1R7 ⁽¹⁾⁽²⁾ with BM Lifecycle Extension for z/OS V1.7	No	Yes
Linux on System z ⁽²⁾ , Red Hat RHEL 4, & Novell SUSE SLES 9	Yes	Yes
Linux on System z ⁽²⁾ , Red Hat RHEL 5, & Novell SUSE SLES 10	No	Yes
z/VM V5R2 ⁽³⁾ , 3 ⁽³⁾ and 4	No*	Yes
z/VSE V3R1 ⁽²⁾⁽⁴⁾	Yes	No
z/VSE V4R1 ⁽²⁾⁽⁵⁾ and 2 ⁽⁵⁾	No	Yes
z/TPF V1R1	No	Yes
TPF V4R1 (ESA mode only)	Yes	No

- z/OS V1.7 support on the z10 BC™ requires the Lifecycle Extension for z/OS V1.7, 5637-A01. The Lifecycle Extension for z/OS R1.7 + zIIP Web Deliverable required for z10 to enable HiperDispatch on z10 (does not require a zIIP). z/OS V1.7 support was withdrawn September 30, 2008. The Lifecycle Extension for z/OS V1.7 (5637-A01) makes fee-based corrective service for z/OS V1.7 available through September 2009. With this Lifecycle Extension, z/OS V1.7 supports the z10 BC server. Certain functions and features of the z10 BC server require later releases of z/OS. For a complete list of software support, see the PSP buckets and the Software Requirements section of the System z10 BC announcement letter, dated October 21, 2008.
- Compatibility Support for listed releases. Compatibility support allows OS to IPL and operate on z10 BC
- Requires Compatibility Support which allows z/VM to IPL and operate on the z10 providing IBM System z9® functionality for the base OS and Guests. *z/VM supports 31-bit and 64-bit guests
- z/VSE V3 operates in 31-bit mode only. It does not implement z/Architecture, and specifically does not implement 64-bit mode capabilities. z/VSE is designed to exploit select features of IBM System z10, System z9, and IBM eServer™ zSeries® hardware.
- z/VSE V4 is designed to exploit 64-bit real memory addressing, but will not support 64-bit virtual memory addressing

Note: Refer to the z/OS, z/VM, z/VSE subsets of the 2098DEVICE Preventive Planning (PSP) bucket prior to installing a z10 BC

Everyday the IT system needs to be available to users – customers that need access to the company Web site, line of business personnel that need access to the system, application development that is constantly keeping the environment current, and the IT staff that is operating and maintaining the environment. If applications are not consistently available, the business can suffer.

The z10 EC continues our commitment to deliver improvements in hardware Reliability, Availability and Serviceability (RAS) with every new System z server. They include micro-code driver enhancements, dynamic segment sparing for memory as well as the fixed HSA. The z10 EC is a server that can help keep applications up and running in the event of planned or unplanned disruptions to the system.

IBM System z servers stand alone against competition and have stood the test of time with our business resiliency solutions. Our coupling solutions with Parallel Sysplex technology allows for greater scalability and availability. The InfiniBand Coupling Links on the z10 EC provides a high speed solution to the 10 meter limitation of ICB-4 since they will be available in lengths up to 150 meters.

What the z10 EC provides over its predecessors are improvements in the processor granularity offerings, more options for specialty engines, security enhancements, additional high availability characteristics, Concurrent Driver Upgrade (CDU) improvements, enhanced networking and on demand offerings. The z10 EC provides our IBM customers an option for continued growth, continuity, and upgradeability.

The IBM System z10 EC builds upon the structure introduced on the IBM System z9 EC – scalability and z/Architecture. The System z10 EC expands upon a key attribute of the platform – availability – to help ensure a resilient infrastructure designed to satisfy the demands

of your business. With the potential for increased performance and capacity, you have an opportunity to continue to consolidate diverse applications on a single platform. The z10 EC is designed to provide up to 1.7 times the total system capacity than the z9 EC, and has up to triple the available memory. The maximum number of Processor Units (PUs) has grown from 54 to 64, and memory has increased from 128 GB per book and 512 GB per system to 384 GB per book and 1.5 TB per system.

The z10 EC will continue to use the Cargo cage for its I/O, supporting up to 960 Channels on the Model E12 (64 I/O features) and up to 1,024 (84 I/O features) on the Models E26, E40, E56 and E64.

HiperDispatch helps provide increased scalability and performance of higher n-way and multi-book z10 EC systems by improving the way workload is dispatched across the server. HiperDispatch accomplishes this by recognizing the physical processor where the work was started and then dispatching subsequent work to the same physical processor. This intelligent dispatching helps reduce the movement of cache and data and is designed to improve CPU time and performance. HiperDispatch is available only with new z10 EC PR/SM and z/OS functions.

Processor Units (cores) defined as Internal Coupling Facilities (ICFs), Integrated Facility for Linux (IFLs), System z10 Application Assist Processor (zAAPs) and System z10 Integrated Information Processor (zIIPs) are no longer grouped together in one pool as on the z990, but are grouped together in their own pool, where they can be managed separately. The separation significantly simplifies capacity planning and management for LPAR and can have an effect on weight management since CP weights and zAAP and zIIP weights can now be managed separately. Capacity BackUp (CBU) features are available for IFLs, ICFs, zAAPs and zIIPs.

For LAN connectivity, z10 EC provides a OSA-Express3 2-port 10 Gigabit Ethernet (GbE) Long Reach feature along with the OSA-Express3 Gigabit Ethernet SX and LX with four ports per features. The z10 EC continues to support OSA-Express2 1000BASE-T and GbE Ethernet features, and supports IP version 6 (IPv6) on HiperSockets. OSA-Express2 OSN (OSA for NCP) is also available on System z10 EC to support the Channel Data Link Control (CDLC) protocol, providing direct access from the host operating system images to the Communication Controller for Linux on the z10 EC, z10 BC, z9 EC and z9 (CCL) using OSA-Express3 or OSA-Express2 to help eliminate the requirement for external hardware for communications.

Additional channel and networking improvements include support for Layer 2 and Layer 3 traffic, FCP management facility for z/VM and Linux for System z, FCP security improvements, and Linux support for HiperSockets IPv6. STP enhancements include the additional support for NTP clients and STP over InfiniBand links.

Like the System z9 EC, the z10 EC offers a configurable Crypto Express2 feature, with PCI-X adapters that can be individually configured as a secure coprocessor or an accelerator for SSL, the TKE workstation with optional Smart Card Reader, and provides the following CP Assist for Cryptographic Function (CPACF):

- *DES, TDES, AES-128, AES-192, AES-256*
- *SHA-1, SHA-224, SHA-256, SHA-384, SHA-512*
- *Pseudo Random Number Generation (PRNG)*

z10 EC is designed to deliver the industry leading Reliability, Availability and Serviceability (RAS) customers expect from System z servers. RAS is designed to reduce all sources of outages by reducing unscheduled, scheduled and planned outages. Planned outages are further designed to be reduced by reducing preplanning requirements.

z10 EC preplanning improvements are designed to avoid planned outages and include:

- *Flexible Customer Initiated Upgrades*
- *Enhanced Driver Maintenance*
 - *Multiple “from” sync point support*
- *Reduce Pre-planning to avoid Power-On-Reset*
 - *16 GB for HSA*
 - *Dynamic I/O enabled by default*
 - *Add Logical Channel Subsystems (LCSS)*
 - *Change LCSS Subchannel Sets*
 - *Add/delete Logical partitions*
- *Designed to eliminate a logical partition deactivate/activate/IPL*
 - *Dynamic Change to Logical Processor Definition – z/VM 5.3*
 - *Dynamic Change to Logical Cryptographic Coprocessor Definition – z/OS ICSF*

Additionally, several service enhancements have also been designed to avoid scheduled outages and include concurrent firmware fixes, concurrent driver upgrades, concurrent parts replacement, and concurrent hardware upgrades. Exclusive to the z10 EC is the ability to hot swap ICB-4 and InfiniBand hub cards.

Enterprises with IBM System z9 EC and IBM z990 may upgrade to any z10 Enterprise Class model. Model upgrades within the z10 EC are concurrent with the exception of the E64, which is disruptive. If you desire a consolidation platform for your mainframe and Linux capable applications, you can add capacity and even expand your current application workloads in a cost-effective manner. If your traditional and new applications are growing, you may find the z10 EC a good fit with its base qualities of service and its specialty processors designed for assisting with new workloads. Value is leveraged with improved hardware price/performance and System z10 EC software pricing strategies.

The z10 EC processor introduces IBM System z10 Enterprise Class with Quad Core technology, advanced pipeline design and enhanced performance on CPU intensive workloads. The z10 EC is specifically designed and optimized for full z/Architecture compatibility. New features enhance enterprise data serving performance, industry leading virtualization capabilities, energy efficiency at system and data center levels. The z10 EC is designed to further extend and integrate key platform characteristics such as dynamic flexible partitioning and resource management in mixed and unpredictable workload environments, providing scalability, high availability and Qualities of Service (QoS) to emerging applications such as WebSphere, Java and Linux.

With the logical partition (LPAR) group capacity limit on z10 EC, z10 BC, z9 EC and z9 BC, you can now specify LPAR group capacity limits allowing you to define each LPAR with its own capacity and one or more groups of LPARs on a server. This is designed to allow z/OS to manage the groups in such a way that the sum of the LPARs' CPU utilization within a group will not exceed the group's defined capacity. Each LPAR in a group can still optionally continue to define an individual LPAR capacity limit.

The z10 EC has five models with a total of 100 capacity settings available as new build systems and as upgrades from the z9 EC and z990.

The five z10 EC models are designed with a multi-book system structure that provides up to 64 Processor Units (PUs) that can be characterized as either Central Processors (CPs), IFLs, ICFs, zAAPs or zIIPs.

Some of the significant enhancements in the z10 EC that help bring improved performance, availability and function to the platform have been identified. The following sections highlight the functions and features of the z10 EC.

z10 EC Design and Technology

The System z10 EC is designed to provide balanced system performance. From processor storage to the system's I/O and network channels, end-to-end bandwidth is provided and designed to deliver data where and when it is needed.

The processor subsystem is comprised of one to four books connected via a point-to-point SMP network. The change to a point-to-point connectivity eliminates the need for the jumper book, as had been used on the System z9 and z990 systems. The z10 EC design provides growth paths up to a 64 engine system where each of the 64 PUs has full access to all system resources, specifically memory and I/O.

Each book is comprised of a Multi-Chip Module (MCM), memory cards and I/O fanout cards. The MCMs, which measure approximately 96 x 96 millimeters, contain the Processor Unit (PU) chips, the "SCD" and "SCC" chips of z990 and z9 have been replaced by a single "SC" chip which includes both the L2 cache and the SMP fabric ("storage controller") functions. There are two SC chips on each MCM, each of which is connected to all five CP chips on that MCM. The MCM contain 103 glass ceramic layers to provide interconnection between the chips and the off-module environment. Four models (E12, E26, E40 and E56) have 17 PUs per book, and the high capacity z10 EC Model E64 has one 17 PU book and three 20 PU books. Each PU measures 21.973 mm x 21.1658 mm and has an L1 cache divided into a 64 KB cache for instructions and a 128 KB cache for data. Each PU also has an L1.5 cache. This cache is 3 MB in size. Each L1 cache has a Translation Look-aside Buffer (TLB) of 512 entries associated with it. The PU, which uses a high-frequency z/Architecture microprocessor core, is built on CMOS 11S chip technology and has a cycle time of approximately 0.23 nanoseconds.

The design of the MCM technology on the z10 EC provides the flexibility to configure the PUs for different uses; there are two spares and up to 11 System Assist Processors (SAPs) standard per system. The remaining inactive PUs on each installed MCM are available to be characterized as either CPs, ICF processors for Coupling Facility applications, or IFLs for Linux applications and z/VM hosting Linux as a guest, System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processors (zIIPs) or as optional SAPs and provide you with tremendous flexibility in establishing the best system for running applications. Each model of the z10 EC must always be ordered with at least one CP, IFL or ICF.

Each book can support from the 16 GB minimum memory, up to 384 GB and up to 1.5 TB per system. 16 GB of the total memory is delivered and reserved for the fixed Hardware Systems Area (HSA). There are up to 48 IFB links per system at 6 GBps each.

The z10 EC supports a combination of Memory Bus Adapter (MBA) and Host Channel Adapter (HCA) fanout cards. New MBA fanout cards are used exclusively for ICB-4. New ICB-4 cables are needed for z10 EC and are only available on models E12, E26, E40 and E56. The E64 model may not have ICBs. The InfiniBand Multiplexer (IFB-MP) card replaces the Self-Timed Interconnect Multiplexer (STI-MP) card. There are two types of HCA fanout cards: HCA2-C is copper and is always used to connect to I/O (IFB-MP card) and the HCA2-O which is optical and used for customer InfiniBand coupling.

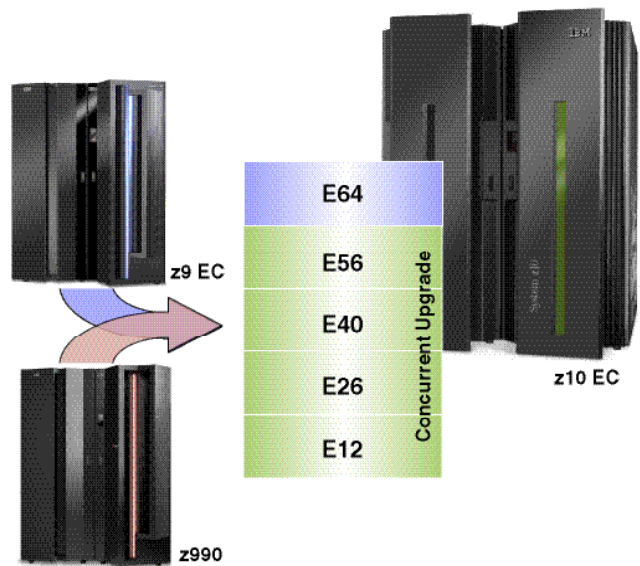
Data transfers are direct between books via the level 2 cache chip in each MCM. Level 2 Cache is shared by all PU chips on the MCM. PR/SM provides the ability to configure and operate as many as 60 Logical Partitions which may be assigned processors, memory and I/O resources from any of the available books.

z10 EC Model

The z10 EC has been designed to offer high performance and efficient I/O structure. All z10 EC models ship with two frames: an A-Frame and a Z-Frame, which together support the installation of up to three I/O cages. The z10 EC will continue to use the Cargo cage for its I/O, supporting up to 960 ESCON® and 256 FICON channels on the Model E12 (64 I/O features) and up to 1,024 ESCON and 336 FICON channels (84 I/O features) on the Models E26, E40, E56 and E64.

To increase the I/O device addressing capability, the I/O subsystem provides support for multiple subchannels sets (MSS), which are designed to allow improved device connectivity for Parallel Access Volumes (PAVs). To support the highly scalable multi-book system design, the z10 EC I/O subsystem uses the Logical Channel Subsystem (LCSS) which provides the capability to install up to 1024 CHPIDs across three I/O cages (256 per operating system image). The Parallel Sysplex Coupling Link architecture and technology continues to support high speed links providing efficient transmission between the Coupling Facility and z/OS systems. HiperSockets provides high-speed capability to communicate among virtual servers and logical partitions. HiperSockets is now improved with the IP version 6 (IPv6) support; this is based on high-speed TCP/IP memory speed transfers and provides value in allowing applications running in one partition to communicate with applications running in another without dependency on an external network. Industry standard and openness are design objectives for I/O in System z10 EC.

The z10 EC has five models offering between 1 to 64 processor units (PUs), which can be configured to provide a highly scalable solution designed to meet the needs of both high transaction processing applications and On Demand Business. Four models (E12, E26, E40 and E56) have 17 PUs per book, and the high capacity z10 EC Model E64 has one 17 PU book and three 20 PU books. The PUs can be characterized as either CPs, IFLs, ICFs, zAAPs or zIIPs. An easy-to-enable ability to “turn off” CPs or IFLs is available on z10 EC, allowing you to purchase capacity for future use with minimal or no impact on software billing. An MES feature will enable the “turned off” CPs or IFLs for use where you require the increased capacity. There are a wide range of upgrade options available in getting to and within the z10 EC.



The z10 EC hardware model numbers (E12, E26, E40, E56 and E64) on their own do not indicate the number of PUs which are being used as CPs. For software billing purposes only, there will be a Capacity Identifier associated with the number of PUs that are characterized as CPs. This

number will be reported by the Store System Information (STSI) instruction for software billing purposes only. There is no affinity between the hardware model and the number of CPs. For example, it is possible to have a Model E26 which has 13 PUs characterized as CPs, so for software billing purposes, the STSI instruction would report 713.

z10 EC model upgrades

There are full upgrades within the z10 EC models and upgrades from any z9 EC or z990 to any z10 EC. Upgrade of z10 EC Models E12, E26, E40 and E56 to the E64 is disruptive. When upgrading to z10 EC Model E64, unlike the z9 EC, the first book is retained. There are no direct upgrades from the z9 BC or IBM eServer zSeries 900 (z900), or previous generation IBM eServer zSeries.

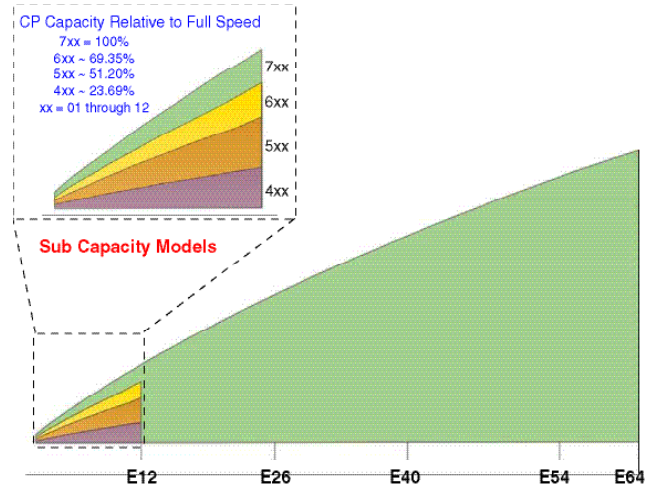
IBM is increasing the number of sub-capacity engines on the z10 EC. A total of 36 sub-capacity settings are available on any hardware model for 1-12 CPs. Models with 13 CPs or greater must be full capacity.

For the z10 EC models with 1-12 CPs, there are four capacity settings per engine for central processors (CPs). The entry point (Model 401) is approximately 23.69% of a full speed CP (Model 701). All specialty engines continue to run at full speed. Sub-capacity processors have availability of z10 EC features/functions and any-to-any upgradeability is available within the sub-capacity matrix. All CPs must be the same capacity setting size within one z10 EC.

z10 EC Model Capacity Identifiers:

- 700, 401 to 412, 501 to 512, 601 to 612 and 701 to 764
- Capacity setting 700 does not have any CP engines
- Nxx, where n = the capacity setting of the engine, and xx = the number of PU characterized as CPs in the CEC
- Once xx exceeds 12, then all CP engines are full capacity

z10 EC Base and Sub-capacity Offerings



- The z10 EC has 36 additional capacity settings at the low end
- Available on ANY H/W Model for 1 to 12 CPs. Models with 13 CPs or greater have to be full capacity
- All CPs must be the same capacity within the z10 EC
- All specialty engines run at full capacity. The one for one entitlement to purchase one zAAP or one zIIP for each CP purchased is the same for CPs of any capacity.
- Only 12 CPs can have granular capacity, other PUs must be CBU or characterized as specialty engines

z10 EC Performance

The performance design of the z/Architecture can enable the server to support a new standard of performance for applications through expanding upon a balanced system approach. As CMOS technology has been enhanced to support not only additional processing power, but also more PUs, the entire server is modified to support the increase in processing power. The I/O subsystem supports a greater amount of bandwidth than previous generations through internal changes, providing for larger and faster volume of data movement into and out of the server. Support of larger amounts of data within the server required improved management of storage configurations, made available through integration of the operating system and hardware support of 64-bit addressing. The combined balanced system design allows for increases in performance across a broad spectrum of work.

Large System Performance Reference

IBM's Large Systems Performance Reference (LSPR) method is designed to provide comprehensive z/Architecture processor capacity ratios for different configurations of Central Processors (CPs) across a wide variety of system control programs and workload environments. For z10 EC, z/Architecture processor capacity identifier is defined with a (7XX) notation, where XX is the number of installed CPs.

Based on using an LSPR mixed workload, the performance of the z10 EC (2097) 701 is expected to be up to 1.62 times that of the z9 EC (2094) 701.

The LSPR contains the Internal Throughput Rate Ratios (ITRRs) for the z10 EC and the previous-generation zSeries processor families based upon measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user

may experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, and the workload processed.

LSPR workloads have been updated to reflect more closely your current and growth workloads. The classification Java Batch (CB-J) has been replaced with a new classification for Java Batch called ODE-B. The remainder of the LSPR workloads are the same as those used for the z9 EC LSPR. The typical LPAR configuration table is used to establish single-number-metrics such as MIPS and MSUs. The z10 EC LSPR will rate all z/Architecture processors running in LPAR mode, 64-bit mode, and assumes that HiperDispatch is enabled.

For more detailed performance information, consult the Large Systems Performance Reference (LSPR) available at: <http://www.ibm.com/servers/eserver/zseries/lSpr/>.

CPU Measurement Facility

The CPU Measurement Facility is a hardware facility which consists of counters and samples. The facility provides a means to collect run-time data for software performance tuning. The detailed architecture information for this facility can be found in the System z10 Library in Resource Link™.

z10 EC I/O Subsystem

The z10 EC contains an I/O subsystem infrastructure which uses an I/O cage that provides 28 I/O slots and the ability to have one to three I/O cages delivering a total of 84 I/O slots. ESCON, FICON Express4, FICON Express2, FICON Express, OSA-Express3, OSA-Express2, and Crypto Express2 features plug into the z10 EC I/O cage along with any ISC-3s and InfiniBand Multiplexer (IFB-MP) cards. All I/O features and their support cards can be hot-plugged in the I/O cage. Installation of an I/O cage remains a disruptive MES, so the Plan Ahead feature remains an important consideration when ordering a z10 EC system. Each model ships with one I/O cage as standard in the A-Frame (the A-Frame also contains the Central Electronic Complex [CEC] cage where the books reside) and any additional I/O cages are installed in the Z-Frame. Each IFB-MP has a bandwidth up to 6 GigaBytes per second (GB/sec) for I/O domains and MBA fanout cards provide 2.0 GB/sec for ICB-4s.

The z10 EC continues to support all of the features announced with the System z9 EC such as:

- *Logical Channel Subsystems (LCSSs) and support for up to 60 logical partitions*
- *Increased number of Subchannels (63.75k)*
- *Multiple Subchannel Sets (MSS)*
- *Redundant I/O Interconnect*
- *Physical Channel IDs (PCHIDs)*
- *System Initiated CHPID Reconfiguration*
- *Logical Channel SubSystem (LCSS) Spanning*

System I/O Configuration Analyzer

Today the information needed to manage a system's I/O configuration has to be obtained from many separate applications. The System's I/O Configuration Analyzer (SIOA) tool is a SE/HMC-based tool that will allow the

system hardware administrator access to the information from these many sources in one place. This will make it much easier to manage I/O configurations, particularly across multiple CPCs. The SIOA is a "view-only" tool. It does not offer any options other than viewing options.

First the SIOA tool analyzes the current active IOCDs on the SE. It extracts information about the defined channel, partitions, link addresses and control units. Next the SIOA tool asks the channels for their node ID information. The FICON channels support remote node ID information, so that is also collected from them. The data is then formatted and displayed on five screens:

- 1) *PCHID Control Unit Screen – Shows PCHIDs, CSS, CHPIDs and their control units*
- 2) *PCHID Partition Screen – Shows PCHIDS, CSS, CHPIDs and what partitions they are in*
- 3) *Control Unit Screen – Shows the control units, their PCHIDs and their link addresses in each of the CSS's*
- 4) *Link Load Screen – Shows the Link address and the PCHIDs that use it*
- 5) *Node ID Screen – Shows the Node ID data under the PCHIDs*

The SIOA tool allows the user to sort on various columns and export the data to a USB flash drive for later viewing.

z10 EC Channels and I/O Connectivity

ESCON Channels

The z10 EC supports up to 1,024 ESCON channels. The high density ESCON feature has 16 ports, 15 of which can be activated for customer use. One port is always reserved as a spare which is activated in the event of a failure of one of the other ports. For high availability the initial order of ESCON features will deliver two 16-port ESCON features and the active ports will be distributed across those features.

Fibre Channel Connectivity

The on demand operating environment requires fast data access, continuous data availability, and improved flexibility, all with a lower cost of ownership. The four port FICON Express4 and FICON Express2 features available on the z9 EC continue to be supported on the System z10 EC.

Choose the FICON Express4 features that best meet your business requirements.

To meet the demands of your Storage Area Network (SAN), provide granularity, facilitate redundant paths, and satisfy your infrastructure requirements, there are three features from which to choose.

Feature	FC #	Infrastructure Feature	Ports per Feature
FICON Express4 10KM LX	3321	Single mode fiber	4
FICON Express4 4KM LX	3324	Single mode fiber	4
FICON Express4 SX	3322	Multimode fiber	4

Choose the features that best meet your granularity, fiber optic cabling, and unrepeated distance requirements.

FICON Express4 Channels

The z10 EC supports up to 336 FICON Express4 channels, each one operating at 1, 2 or 4 Gb/sec auto-negotiated. The FICON Express4 features are available in long wavelength (LX) and short wavelength (SX). For customers exploiting LX, there are two options available for unrepeated distances of up to 4 kilometers (2.5 miles) or up to 10 kilometers (6.2 miles). Both LX features use 9 micron single mode fiber optic cables. The SX feature uses 50 or 62.5 micron multimode fiber optic cables. Each FICON Express4 feature has four independent channels (ports) and can be configured to carry native FICON traffic or Fibre Channel (SCSI) traffic. LX and SX cannot be intermixed on a single feature. The receiving devices must correspond to the appropriate LX or SX feature. The maximum number of FICON Express4 features is 84 using three I/O cages.

FICON Express2 Channels

The z10 EC supports carrying forward up to 336 FICON Express2 channels, each one operating at 1 or 2 Gb/sec auto-negotiated. The FICON Express2 features are available in long wavelength (LX) using 9 micron single mode fiber optic cables and short wavelength (SX) using 50 and 62.5 micron multimode fiber optic cables. Each FICON Express2 feature has four independent channels (ports) and each can be configured to carry native FICON traffic or Fibre Channel (SCSI) traffic. LX and SX cannot be intermixed on a single feature. The maximum number of FICON Express2 features is 84, using three I/O cages.

FICON Express Channels

The z10 EC also supports carrying forward FICON Express LX and SX channels from z9 EC and z990 (up to 120 channels) each channel operating at 1 or 2 Gb/sec auto-negotiated. Each FICON Express feature has two independent channels (ports).

The System z10 EC Model E12 is limited to 64 features – any combination of FICON Express4, FICON Express2 and FICON Express LX and SX features.

The FICON Express4, FICON Express2 and FICON Express feature conforms to the Fibre Connection (FICON) architecture and the Fibre Channel (FC) architecture, providing connectivity between any combination of servers, directors, switches, and devices in a Storage Area Network (SAN). Each of the four independent channels (FICON Express only supports two channels per feature) is capable of 1 Gigabit per second (Gb/sec), 2 Gb/sec, or 4 Gb/sec (only FICON Express4 supports 4 Gbps) depending upon the capability of the attached switch or device. The link speed is auto-negotiated, point-to-point, and is transparent to users and applications. Not all switches and devices support 2 or 4 Gb/sec link data rates.

FICON Express4 and FICON Express2 Performance

Your enterprise may benefit from FICON Express4 and FICON Express2 with:

- *Increased data transfer rates (bandwidth)*
- *Improved performance*
- *Increased number of start I/Os*
- *Reduced backup windows*
- *Channel aggregation to help reduce infrastructure costs*

For more information about FICON, visit the IBM Redbooks® Web site at: <http://www.redbooks.ibm.com/> search for SG24-5444. There are also various FICON I/O Connectivity information at: www-03.ibm.com/systems/z/connectivity/.

Concurrent Update

The FICON Express4 SX and LX features may be added to an existing z10 EC concurrently. This concurrent update capability allows you to continue to run workloads through other channels while the new FICON Express4 features are being added. This applies to CHPID types FC and FCP.

Continued Support of Spanned Channels and Logical Partitions

The FICON Express4 and FICON Express2, FICON and FCP (CHPID types FC and FCP) channel types, can be defined as a spanned channel and can be shared among logical partitions within and across LCSSs.

Modes of Operation

There are two modes of operation supported by FICON Express4 and FICON Express2 SX and LX. These modes are configured on a channel-by-channel basis – each of the four channels can be configured in either of two supported modes.

- *Fibre Channel (CHPID type FC), which is native FICON or FICON Channel-to-Channel (server-to-server)*
- *Fibre Channel Protocol (CHPID type FCP), which supports attachment to SCSI devices via Fibre Channel switches or directors in z/VM, z/VSE, and Linux on System z10 environments*

Native FICON Channels

Native FICON channels and devices can help to reduce bandwidth constraints and channel contention to enable easier server consolidation, new application growth, large business intelligence queries and exploitation of On Demand Business.

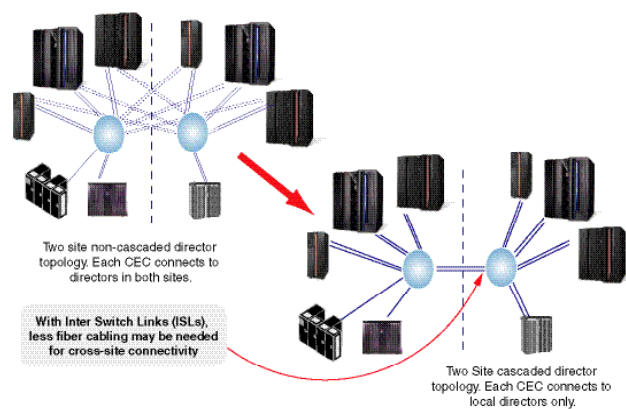
The FICON Express4, FICON Express2 and FICON Express channels support native FICON and FICON Channel-to-Channel (CTC) traffic for attachment to servers, disks, tapes, and printers that comply with the FICON architecture. Native FICON is supported by all of the z10 EC operating systems. Native FICON and FICON CTC are defined as CHPID type FC.

Because the FICON CTC function is included as part of the native FICON (FC) mode of operation, FICON CTC is not limited to intersystem connectivity (as is the case with ESCON), but will support multiple device definitions.

FICON Support for Cascaded Directors

Native FICON (FC) channels support cascaded directors. This support is for a single hop configuration only. Two-director cascading requires a single vendor high integrity fabric. Directors must be from the same vendor since cascaded architecture implementations can be unique. This type of cascaded support is important for disaster recovery and business continuity solutions because it can help provide high availability, extended distance connectivity, and (particularly with the implementation of 2 Gb/sec Inter Switch Links) has the potential for fiber infrastructure cost savings by reducing the number of channels for interconnecting the two sites.

FICON cascaded directors have the added value of high integrity connectivity. Integrity features introduced within the FICON Express channel and the FICON cascaded switch fabric to aid in the detection and reporting of any miscabling actions occurring within the fabric can prevent data from being delivered to the wrong end point.



FCP Channels

z10 EC supports FCP channels, switches and FCP/SCSI disks with full fabric connectivity under Linux on System z and z/VM 5.2 (or later) for Linux as a guest under z/VM, under z/VM 5.2 (or later), and under z/VSE 3.1 for system

usage including install and IPL. Support for FCP devices means that z10 EC servers are capable of attaching to select FCP-attached SCSI devices and may access these devices from Linux on z10 EC and z/VSE. This expanded attachability means that enterprises have more choices for new storage solutions, or may have the ability to use existing storage devices, thus leveraging existing investments and lowering total cost of ownership for their Linux implementations.

The same FICON features used for native FICON channels can be defined to be used for Fibre Channel Protocol (FCP) channels. FCP channels are defined as CHPID type FCP. The 4 Gb/sec capability on the FICON Express4 channel means that 4 Gb/sec link data rates are available for FCP channels as well.

FCP – increased performance for small block sizes

The Fibre Channel Protocol (FCP) Licensed Internal Code has been modified to help provide increased I/O operations per second for small block sizes. With FICON Express4, there may be up to 57,000 I/O operations per second (all reads, all writes, or a mix of reads and writes), an 80% increase compared to System z9. These results are achieved in a laboratory environment using one channel configured as CHPID type FCP with no other processing occurring and do not represent actual field measurements. A significant increase in I/O operations per second for small block sizes can also be expected with FICON Express2.

This FCP performance improvement is transparent to operating systems that support FCP, and applies to all the FICON Express4 and FICON Express2 features when configured as CHPID type FCP, communicating with SCSI devices.

SCSI IPL now a base function

The SCSI Initial Program Load (IPL) enablement feature, first introduced on z990 in October of 2003, is no longer required. The function is now delivered as a part of the server Licensed Internal Code. SCSI IPL allows an IPL of an operating system from an FCP-attached SCSI disk.

FCP Full fabric connectivity

FCP full fabric support means that any number of (single vendor) FCP directors/ switches can be placed between the server and an FCP/SCSI device, thereby allowing many “hops” through a Storage Area Network (SAN) for I/O connectivity. FCP full fabric connectivity enables multiple FCP switches/directors on a fabric to share links and therefore provides improved utilization of inter-site connected resources and infrastructure.

FICON and FCP for connectivity to disk, tape, and printers

High Performance FICON – improvement in performance and RAS

Enhancements have been made to the z/Architecture and the FICON interface architecture to deliver optimizations for online transaction processing (OLTP) workloads. When exploited by the FICON channel, the z/OS operating system, and the control unit, High Performance FICON for System z (zHPF) is designed to help reduce overhead and improve performance.

Additionally, the changes to the architectures offer end-to-end system enhancements to improve reliability, availability, and serviceability (RAS).

zHPF channel programs can be exploited by the OLTP I/O workloads – DB2, VSAM, PDSE, and zFS – which transfer small blocks of fixed size data (4K blocks). zHPF implementation by the IBM System Storage DS8000™ is exclusively for I/Os that transfer less than a single track of data.

The maximum number of I/Os is designed to be improved up to 100% for small data transfers that can exploit zHPF. Realistic production workloads with a mix of data transfer sizes can see up to 30 to 70% of FICON I/Os utilizing zHPF resulting in up to a 10 to 30% savings in channel utilization. Sequential I/Os transferring less than a single track size (for example, 12x4k bytes/IO) may also benefit.

The FICON Express4 and FICON Express2 features will support both the existing FICON protocol and the zHPF protocol concurrently in the server Licensed Internal Code. High performance FICON is supported by z/OS for DB2, VSAM, PDSE, and zFS applications. zHPF applies to all FICON Express4 and FICON Express2 features (CHPID type FC) and is exclusive to System z10. Exploitation is required by the control unit.

IBM System Storage DS8000 Release 4.1 delivers new capabilities to support High Performance FICON for System z, which can improve FICON I/O throughput on a DS8000 port by up to 100%. The DS8000 series Licensed Machine Code (LMC) level 5.4.2xx.xx (bundle version 64.2.xx.xx), or later, is required.

Platform and name server registration in FICON channel

The FICON channel now provides the same information to the fabric as is commonly provided by open systems, registering with the name server in the attached FICON directors. With this information, your storage area network (SAN) can be more easily and efficiently managed, enhancing your ability to perform problem determination and analysis.

Registration allows other nodes and/or SAN managers to query the name server to determine what is connected to the fabric, what protocols are supported (FICON, FCP) and to gain information about the System z10 using the attributes that are registered. The FICON channel is now designed to perform registration with the fibre channel's Management Service and Directory Service.

It will register:

- *Platform's:*
 - *Worldwide node name (node name for the platform - same for all channels)*
 - *Platform type (host computer)*
 - *Platform name (includes vendor ID, product ID, and vendor specific data from the node descriptor)*
- *Channel's:*
 - *Worldwide port name (WWPN)*
 - *Node port identification (N_PORT ID)*
 - *FC-4 types supported (always 0x1B and additionally 0x1C if any Channel-to-Channel (CTC) control units are defined on that channel)*
 - *Classes of service support by the channel*

Platform registration is a service defined in the Fibre Channel - Generic Services 4 (FC-GS-4) standard (INCITS (ANSI) T11 group).

Platform and name server registration applies to all of the FICON Express4, FICON Express2, and FICON Express features (CHPID type FC). This support is exclusive to System z10 and is transparent to operating systems.

Preplanning and setup of SAN for a System z10 environment

The worldwide port name (WWPN) prediction tool is now available to assist you with preplanning of your Storage Area Network (SAN) environment prior to the installation of your System z10 server.

This standalone tool is designed to allow you to setup your SAN in advance, so that you can be up and running much faster once the server is installed. The tool assigns WWPNs to each virtual Fibre Channel Protocol (FCP) channel/port using the same WWPN assignment algorithms a system uses when assigning WWPNs for channels utilizing N_Port Identifier Virtualization (NPIV).

The tool needs to know the FCP-specific I/O device definitions in the form of a .csv file. This file can either be created manually, or exported from Hardware Configuration Definition/Hardware Configuration Manager (HCD/HCM). The tool will then create the WWPN assignments, which are required to set up your SAN. The tool will also create a binary configuration file that can later on be imported by your system.

The WWPN prediction tool can be downloaded from Resource Link and is applicable to all FICON channels defined as CHPID type FCP (for communication with SCSI devices). Check Preventive Service Planning (PSP) buckets for required maintenance.

<http://www.ibm.com/servers/resourceink/>

Extended distance FICON – improved performance at extended distance

An enhancement to the industry standard FICON architecture (FC-SB-3) helps avoid degradation of performance at extended distances by implementing a new protocol for “persistent” Information Unit (IU) pacing. Control units that exploit the enhancement to the architecture can increase the pacing count (the number of IUs allowed to be in flight from channel to control unit). Extended distance FICON also allows the channel to “remember” the last pacing update for use on subsequent operations to help avoid degradation of performance at the start of each new operation.

Improved IU pacing can help to optimize the utilization of the link, for example help keep a 4 Gbps link fully utilized at 50 km, and allows channel extenders to work at any distance, with performance results similar to that experienced when using emulation.

The requirements for channel extension equipment are simplified with the increased number of commands in flight. This may benefit z/OS Global Mirror (Extended Remote Copy – XRC) applications as the channel extension kit is no longer required to simulate specific channel commands. Simplifying the channel extension requirements may help reduce the total cost of ownership of end-to-end solutions.

Extended distance FICON is transparent to operating systems and applies to all the FICON Express2 and FICON Express4 features carrying native FICON traffic (CHPID type FC). For exploitation, the control unit must support the new IU pacing protocol. The channel will default to current pacing values when operating with control units that cannot exploit extended distance FICON.

Exploitation of extended distance FICON is supported by IBM System Storage DS8000 series Licensed Machine Code (LMC) level 5.3.1xx.xx (bundle version 63.1.xx.xx), or later.

To support extended distance without performance degradation, the buffer credits in the FICON director must be set appropriately. The number of buffer credits required is dependent upon the link data rate (1 Gbps, 2 Gbps, or 4 Gbps), the maximum number of buffer credits supported by the FICON director or control unit, as well as application and workload characteristics. High bandwidth at extended distances is achievable only if enough buffer credits exist to support the link data rate.

FICON Express enhancements for Storage Area Networks

N_Port ID Virtualization

N_Port ID Virtualization is designed to allow for sharing of a single physical FCP channel among multiple operating system images. Virtualization function is currently available for ESCON and FICON channels, and is now available for FCP channels. This function offers improved FCP channel

utilization due to fewer hardware requirements, and can reduce the complexity of physical FCP I/O connectivity.

Program Directed re-IPL

Program Directed re-IPL is designed to enable an operating system to determine how and from where it had been loaded. Further, Program Directed re-IPL may then request that it be reloaded again from the same load device using the same load parameters. In this way, Program Directed re-IPL allows a program running natively in a partition to trigger a re-IPL. This re-IPL is supported for both SCSI and ECKD™ devices. z/VM 5.3 provides support for guest exploitation.

FICON Link Incident Reporting

FICON Link Incident Reporting is designed to allow an operating system image (without operating intervention) to register for link incident reports, which can improve the ability to capture data for link error analysis. The information can be displayed and is saved in the system log.

Serviceability Enhancements

Requests Node Identification Data (RNID) is designed to facilitate the resolution of fiber optic cabling problems. You can now request RNID data for a device attached to a native FICON channel.

Local Area Network (LAN) connectivity –

OSA-Express3 – the newest family of LAN adapters

The third generation of Open Systems Adapter-Express (OSA-Express3) features have been introduced to help reduce latency and overhead, deliver double the port density of OSA-Express2, and provide increased throughput.

Choose the OSA-Express3 features that best meet your business requirements.

To meet the demands of your applications, provide granularity, facilitate redundant paths, and satisfy your infrastructure requirements, there are five features from which to choose. In the 10 GbE environment, Short Reach (SR) is being offered for the first time.

Feature	Infrastructure	Ports per Feature
OSA-Express3 GbE LX	Single mode fiber	4
OSA-Express3 10 GbE LR	Single mode fiber	2
OSA-Express3 GbE SX	Multimode fiber	4
OSA-Express3 10 GbE SR	Multimode fiber	2
OSA-Express3 1000BASE-T	Copper	4

Note that software PTFs or a new release may be required (depending on CHPID type) to support all ports.

OSA-Express3 for reduced latency and improved throughput

To help reduce latency, the OSA-Express3 features now have an Ethernet hardware data router; what was previously done in firmware (packet construction, inspection, and routing) is now performed in hardware. With direct memory access, packets flow directly from host memory to the LAN without firmware intervention. OSA-Express3 is also designed to help reduce the round-trip networking time between systems. Up to a 45% reduction in latency at the TCP/IP application layer has been measured.

The OSA-Express3 features are also designed to improve throughput for standard frames (1492 byte) and jumbo frames (8992 byte) to help satisfy the bandwidth requirements of your applications. Up to a 4x improvement has been measured (compared to OSA-Express2).

The above statements are based on OSA-Express3 performance measurements performed in a laboratory environment on a System z10 and do not represent actual field measurements. Results may vary.

Port density or granularity

The OSA-Express3 features have Peripheral Component Interconnect Express (PCI-E) adapters. The previous table identifies whether the feature has two or four ports for LAN connectivity. Select the density that best meets your business requirements. Doubling the port density on a single feature helps to reduce the number of I/O slots required for high-speed connectivity to the Local Area Network.

The OSA-Express3 10 GbE features support Long Reach (LR) using 9 micron single mode fiber optic cabling and Short Reach (SR) using 50 or 62.5 micron multimode fiber optic cabling. The connector is new; it is now the small form factor, LC Duplex connector. Previously the SC Duplex connector was supported for LR. The LC Duplex connector is common with FICON, ISC-3, and OSA-Express2 Gigabit Ethernet LX and SX.

The OSA-Express3 features are exclusive to System z10.

There are operating system dependencies for exploitation of two ports in OSD mode per PCI-E adapter. Whether it is a 2-port or a 4-port feature, only one of the ports will be visible on a PCI-E adapter if operating system exploitation updates are not installed.

OSA-Express3 Ethernet features – Summary of benefits

OSA-Express3 10 GbE LR (single mode fiber), 10 GbE SR (multimode fiber), GbE LX (single mode fiber), GbE SX (multimode fiber), and 1000BASE-T (copper) are designed for use in high-speed enterprise backbones, for local area network connectivity between campuses, to connect server farms to System z10, and to consolidate file servers

onto System z10. With reduced latency, improved throughput, and up to 96 ports of LAN connectivity, (when all are 4-port features, 24 features per server), you can “do more with less.”

The key benefits of OSA-Express3 compared to OSA-Express2 are:

- *Reduced latency (up to 45% reduction) and increased throughput (up to 4x) for applications*
- *More physical connectivity to service the network and fewer required resources:*
 - *Fewer CHPIDs to define and manage*
 - *Reduction in the number of required I/O slots*
 - *Possible reduction in the number of I/O drawers*
 - *Double the port density of OSA-Express2*
 - *A solution to the requirement for more than 48 LAN ports (now up to 96 ports)*

The OSA-Express3 features are exclusive to System z10.

OSA-Express2 availability

OSA-Express2 Gigabit Ethernet and 1000BASE-T Ethernet continue to be available for ordering, for a limited time, if you are not yet in a position to migrate to the latest release of the operating system for exploitation of two ports per PCI-E adapter and if you are not resource-constrained.

Historical summary: Functions that continue to be supported by OSA-Express3 and OSA-Express2

- *Queued Direct Input/Output (QDIO) – uses memory queues and a signaling protocol to directly exchange data between the OSA microprocessor and the network software for high-speed communication.*
 - *QDIO Layer 2 (Link layer) – for IP (IPv4, IPv6) or non-IP (AppleTalk, DECnet, IPX, NetBIOS, or SNA) workloads. Using this mode the Open Systems Adapter (OSA) is protocol-independent and Layer-3 independent. Packet forwarding decisions are based upon the*

Medium Access Control (MAC) address.

- *QDIO Layer 3 (Network or IP layer) – for IP workloads. Packet forwarding decisions are based upon the IP address. All guests share OSA’s MAC address.*
- *Jumbo frames in QDIO mode (8992 byte frame size) when operating at 1 Gbps (fiber or copper) and 10 Gbps (fiber).*
- *640 TCP/IP stacks per CHPID – for hosting more images.*
- *Large send for IPv4 packets – for TCP/IP traffic and CPU efficiency, offloading the TCP segmentation processing from the host TCP/IP stack to the OSA-Express feature.*
- *Concurrent LIC update – to help minimize the disruption of network traffic during an update; when properly configured, designed to avoid a configuration off or on (applies to CHPID types OSD and OSN).*
- *Multiple Image Facility (MIF) and spanned channels – for sharing OSA among logical channel subsystems*

The OSA-Express3 and OSA-Express2 Ethernet features support the following CHPID types:

CHPID Type	OSA-Express3, OSA-Express2 Features	Purpose/Traffic
OSC	1000BASE-T	OSA-Integrated Console Controller (OSA-ICC) TN3270E, non-SNA DFT, IPL to CPC and LPARs Operating system console operations
OSD	1000BASE-T GbE 10 GbE	Queued Direct Input/Output (QDIO) TCP/IP traffic when Layer 3 Protocol-independent when Layer 2
OSE	1000BASE-T passthru (LCS)	Non-QDIO, SNA/APPN®/HPR and/or TCP/IP
OSN	1000BASE-T GbE	OSA for NCP Supports channel data link control (CDLC)

OSA-Express3 10 GbE

OSA-Express3 10 Gigabit Ethernet LR

The OSA-Express3 10 Gigabit Ethernet (GbE) long reach (LR) feature has two ports. Each port resides on a PCIe adapter and has its own channel path identifier (CHPID).

There are two PCIe adapters per feature. OSA-Express3 10 GbE LR is designed to support attachment to a 10 Gigabits per second (Gbps) Ethernet Local Area Network (LAN) or Ethernet switch capable of 10 Gbps. OSA-Express3 10 GbE LR supports CHPID type OSD exclusively. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

OSA-Express3 10 Gigabit Ethernet SR

The OSA-Express3 10 Gigabit Ethernet (GbE) short reach (SR) feature has two ports. Each port resides on a PCIe adapter and has its own channel path identifier (CHPID). There are two PCIe adapters per feature. OSA-Express3 10 GbE SR is designed to support attachment to a 10 Gigabits per second (Gbps) Ethernet Local Area Network (LAN) or Ethernet switch capable of 10 Gbps. OSA-Express3 10 GbE SR supports CHPID type OSD exclusively. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

OSA-Express3 Gigabit Ethernet LX

The OSA-Express3 Gigabit Ethernet (GbE) long wavelength (LX) feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port supports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN). OSA-Express3 GbE LX supports CHPID types OSD and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

OSA-Express3 Gigabit Ethernet SX

The OSA-Express3 Gigabit Ethernet (GbE) short wavelength (SX) feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port sup-

ports attachment to a one Gigabit per second (Gbps) Ethernet Local Area Network (LAN). OSA-Express3 GbE SX supports CHPID types OSD and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

Four-port exploitation on OSA-Express3 GbE SX and LX

For the operating system to recognize all four ports on an OSA-Express3 Gigabit Ethernet feature, a new release and/or PTF is required. If software updates are not applied, only two of the four ports will be “visible” to the operating system.

Activating all four ports on an OSA-Express3 feature provides you with more physical connectivity to service the network and reduces the number of required resources (I/O slots, I/O cages, fewer CHPIDs to define and manage).

Four-port exploitation is supported by z/OS, z/VM, z/VSE, z/TPF, and Linux on System z.

OSA-Express3 1000BASE-T Ethernet

The OSA-Express3 1000BASE-T Ethernet feature has four ports. Two ports reside on a PCIe adapter and share a channel path identifier (CHPID). There are two PCIe adapters per feature. Each port supports attachment to either a 10BASE-T (10 Mbps), 100BASE-TX (100 Mbps), or 1000BASE-T (1000 Mbps or 1 Gbps) Ethernet Local Area Network (LAN). The feature supports auto-negotiation and automatically adjusts to 10, 100, or 1000 Mbps, depending upon the LAN. When the feature is set to autonegotiate, the target device must also be set to autonegotiate. The feature supports the following settings: 10 Mbps half or full duplex, 100 Mbps half or full duplex, 1000 Mbps (1 Gbps) full duplex. OSA-Express3 1000BASE-T Ethernet supports CHPID types OSC, OSD, OSE, and OSN. It can be defined as a spanned channel and can be shared among LPARs within and across LCSSs.

When configured at 1 Gbps, the 1000BASE-T Ethernet feature operates in full duplex mode only and supports jumbo frames when in QDIO mode (CHPID type OSD).

OSA-Express QDIO data connection isolation for the z/VM environment

Multi-tier security zones are fast becoming the network configuration standard for new workloads. Therefore, it is essential for workloads (servers and clients) hosted in a virtualized environment (shared resources) to be protected from intrusion or exposure of data and processes from other workloads.

With Queued Direct Input/Output (QDIO) data connection isolation you:

- *Have the ability to adhere to security and HIPAA-security guidelines and regulations for network isolation between the operating system instances sharing physical network connectivity*
- *Can establish security zone boundaries that have been defined by your network administrators*
- *Have a mechanism to isolate a QDIO data connection (on an OSA port), ensuring all internal OSA routing between the isolated QDIO data connections and all other sharing QDIO data connections is disabled. In this state, only external communications to and from the isolated QDIO data connection are allowed. If you choose to deploy an external firewall to control the access between hosts on an isolated virtual switch and sharing LPARs then an external firewall needs to be configured and each individual host and or LPAR must have a route added to their TCP/IP stack to forward local traffic to the firewall.*

Internal “routing” can be disabled on a per QDIO connection basis. This support does not affect the ability to share an OSA-Express port. Sharing occurs as it does today, but the ability to communicate between sharing QDIO data connections may be restricted through the use of this support. You decide whether an operating system’s or z/VM’s

Virtual Switch OSA-Express QDIO connection is to be non-isolated (default) or isolated.

QDIO data connection isolation applies to the device statement defined at the operating system level. While an OSA-Express CHPID may be shared by an operating system, the data device is not shared.

QDIO data connection isolation applies to the z/VM 5.3 and 5.4 with PTFs environment and to all of the OSA-Express3 and OSA-Express2 features (CHPID type OSD) on System z10 and to the OSA-Express2 features on System z9.

Network Traffic Analyzer

With the large volume and complexity of today’s network traffic, the z10 EC offers systems programmers and network administrators the ability to more easily solve network problems. With the introduction of the OSA-Express Network Traffic Analyzer and QDIO Diagnostic Synchronization on the System z and available on the z10 EC, customers will have the ability to capture trace/trap data and forward it to z/OS 1.8 tools for easier problem determination and resolution.

This function is designed to allow the operating system to control the sniffer trace for the LAN and capture the records into host memory and storage (file systems), using existing host operating system tools to format, edit, and process the sniffer records.

OSA-Express Network Traffic Analyzer is exclusive to the z10 EC, z10 BC, z9 EC and z9 BC, and is applicable to the OSA-Express3 and OSA-Express2 features when configured as CHPID type OSD (QDIO), and is supported by z/OS.

Dynamic LAN idle for z/OS

Dynamic LAN idle is designed to reduce latency and improve network performance by dynamically adjusting the inbound blocking algorithm. When enabled, the z/OS TCP/IP stack is designed to adjust the inbound blocking algorithm to best match the application requirements.

For latency sensitive applications, the blocking algorithm is modified to be “latency sensitive.” For streaming (throughput sensitive) applications, the blocking algorithm is adjusted to maximize throughput. The z/OS TCP/IP stack can dynamically detect the application requirements, making the necessary adjustments to the blocking algorithm. The monitoring of the application and the blocking algorithm adjustments are made in real-time, dynamically adjusting the application’s LAN performance.

System administrators can authorize the z/OS TCP/IP stack to enable a dynamic setting, which was previously a static setting. The z/OS TCP/IP stack is able to help determine the best setting for the current running application, based on system configuration, inbound workload volume, CPU utilization, and traffic patterns.

Link aggregation for z/VM in Layer 2 mode

z/VM Virtual Switch-controlled (VSWITCH-controlled) link aggregation (IEEE 802.3ad) allows you to dedicate an OSA-Express2 (or OSA-Express3) port to the z/VM operating system when the port is participating in an aggregated group when configured in Layer 2 mode. Link aggregation (trunking) is designed to allow you to combine multiple physical OSA-Express3 and OSA-Express2 ports (of the same type for example 1GbE or 10GbE) into a single logical link for increased throughput and for non-disruptive failover in the event that a port becomes unavailable.

- *Aggregated link viewed as one logical trunk and containing all of the Virtual LANs (VLANs) required by the LAN segment*
- *Load balance communications across several links in a trunk to prevent a single link from being overrun*
- *Link aggregation between a VSWITCH and the physical network switch*
- *Point-to-point connections*
- *Up to eight OSA-Express3 or OSA-Express2 ports in one aggregated link*

- *Ability to dynamically add/remove OSA ports for “on demand” bandwidth*
- *Full-duplex mode (send and receive)*
- *Target links for aggregation must be of the same type (for example, Gigabit Ethernet to Gigabit Ethernet)*

The Open Systems Adapter/Support Facility (OSA/SF) will provide status information on an OSA port – its “shared” or “exclusive use” state. OSA/SF is an integrated component of z/VM.

Link aggregation is exclusive to System z10 and System z9, is applicable to the OSA-Express3 and OSA-Express2 features in Layer 2 mode when configured as CHPID type OSD (QDIO), and is supported by z/VM 5.3 and later.

Layer 2 transport mode: When would it be used?

If you have an environment with an abundance of Linux images in a guest LAN environment, or you need to define router guests to provide the connection between these guest LANs and the OSA-Express3 features, then using the Layer 2 transport mode may be the solution. If you have Internet-network Packet Exchange (IPX), NetBIOS, and SNA protocols, in addition to Internet Protocol Version 4 (IPv4) and IPv6, use of Layer 2 could provide “protocol independence.”

The OSA-Express3 features have the capability to perform like Layer 2 type devices, providing the capability of being protocol- or Layer-3-independent (that is, not IP-only). With the Layer 2 interface, packet forwarding decisions are based upon Link Layer (Layer 2) information, instead of Network Layer (Layer 3) information. Each operating system attached to the Layer 2 interface uses its own MAC address. This means the traffic can be IPX, NetBIOS, SNA, IPv4, or IPv6.

An OSA-Express3 feature can filter inbound datagrams by Virtual Local Area Network identification (VLAN ID, IEEE 802.1q), and/or the Ethernet destination MAC address. Filtering can reduce the amount of inbound traffic being processed by the operating system, reducing CPU utilization.

Layer 2 transport mode is supported by z/VM and Linux on System z.

OSA Layer 3 Virtual MAC for z/OS

To simplify the infrastructure and to facilitate load balancing when an LPAR is sharing the same OSA Media Access Control (MAC) address with another LPAR, each operating system instance can now have its own unique “logical” or “virtual” MAC (VMAC) address. All IP addresses associated with a TCP/IP stack are accessible using their own VMAC address, instead of sharing the MAC address of an OSA port. This applies to Layer 3 mode and to an OSA port shared among Logical Channel Subsystems.

This support is designed to:

- *Improve IP workload balancing*
- *Dedicate a Layer 3 VMAC to a single TCP/IP stack*
- *Remove the dependency on Generic Routing Encapsulation (GRE) tunnels*
- *Improve outbound routing*
- *Simplify configuration setup*
- *Allow WebSphere Application Server content-based routing to work with z/OS in an IPv6 network*
- *Allow z/OS to use a “standard” interface ID for IPv6 addresses*
- *Remove the need for PRIROUTER/SECROUTER function in z/OS*

OSA Layer 3 VMAC for z/OS is exclusive to System z, and is applicable to OSA-Express3 and OSA-Express2 features when configured as CHPID type OSD (QDIO).

Direct Memory Access (DMA)

OSA-Express3 and the operating systems share a common storage area for memory-to-memory communication, reducing system overhead and improving performance. There are no read or write channel programs for data exchange. For write processing, no I/O interrupts have to be handled. For read processing, the number of I/O interrupts is minimized.

Hardware data router

With OSA-Express3, much of what was previously done in firmware (packet construction, inspection, and routing) is now performed in hardware. This allows packets to flow directly from host memory to the LAN without firmware intervention.

With the hardware data router, the “store and forward” technique is no longer used, which enables true direct memory access, a direct host memory-to-LAN flow, returning CPU cycles for application use.

This avoids a “hop” and is designed to reduce latency and to increase throughput for standard frames (1492 byte) and jumbo frames (8992 byte).

IBM Communication Controller for Linux (CCL)

CCL is designed to help eliminate hardware dependencies, such as 3745/3746 Communication Controllers, ESCON channels, and Token Ring LANs, by providing a software solution that allows the Network Control Program (NCP) to be run in Linux on System z freeing up valuable data center floor space.

CCL helps preserve mission critical SNA functions, such as SNI, and z/OS applications workloads which depend upon these functions, allowing you to collapse SNA inside a z10 EC while exploiting and leveraging IP.

The OSA-Express3 and OSA-Express2 GbE and 1000BASE-T Ethernet features provide support for CCL. This support is designed to require no changes to operating systems (does require a PTF to support CHPID type OSN) and also allows TPF to exploit CCL. Supported by z/VM for Linux and z/TPF guest environments.

OSA-Express3 and OSA-Express2 OSN (OSA for NCP)

OSA-Express for Network Control Program (NCP), Channel path identifier (CHPID) type OSN, is now available for use with the OSA-Express3 GbE features as well as the OSA-Express3 1000BASE-T Ethernet features.

OSA-Express for NCP, supporting the channel data link control (CDLC) protocol, provides connectivity between System z operating systems and IBM Communication Controller for Linux (CCL). CCL allows you to keep your business data and applications on the mainframe operating systems while moving NCP functions to Linux on System z.

CCL provides a foundation to help enterprises simplify their network infrastructure while supporting traditional Systems Network Architecture (SNA) functions such as SNA Network Interconnect (SNI).

Communication Controller for Linux on System z is the solution for companies that want to help improve network availability by replacing token-ring networks and ESCON channels with an Ethernet network and integrated LAN adapters on System z10, OSA-Express3 or OSA-Express2 GbE or 1000BASE-T.

OSA-Express for NCP is supported in the z/OS, z/VM, z/VSE, TPF, z/TPF, and Linux on System z environments.

OSA Integrated Console Controller

The OSA-Express Integrated Console Controller (OSA-ICC) support is a no-charge function included in Licensed Internal Code (LIC) on z10 EC, z10 BC, z9 EC, z9 BC, z990, and z890 servers. It is available via the OSA-Express3, OSA-Express2 and OSA-Express 1000BASE-T Ethernet features, and supports Ethernet-attached TN3270E consoles.

The OSA-ICC provides a system console function at IPL time and operating systems support for multiple logical partitions. Console support can be used by z/OS, z/OS.e, z/VM, z/VSE, z/TPF, and TPF. The OSA-ICC also supports local non-SNA DFT 3270 and 328x printer emulation for TSO/E, CICS, IMS™, or any other 3270 application that communicates through VTAM®.

With the OSA-Express3 and OSA-Express2 1000BASE-T Ethernet features, the OSA-ICC is configured on a port by port basis, using the Channel Path Identifier (CHPID) type OSC. Each port can support up to 120 console session connections, can be shared among logical partitions using Multiple Image Facility (MIF), and can be spanned across multiple Channel Subsystems (CSSs).

Remove L2/L3 LPAR-to-LPAR Restriction

OSA port sharing between virtual switches can communicate whether the transport mode is the same (Layer 2 to Layer 2) or different (Layer 2 to Layer 3). This enhancement is designed to allow seamless mixing of Layer 2 and Layer 3 traffic, helping to reduce the total cost of networking. Previously, Layer 2 and Layer 3 TCP/IP connections through the same OSA port (CHPID) were unable to communicate with each other LPAR-to-LPAR using the Multiple Image Facility (MIF).

This enhancement is designed to facilitate a migration from Layer 3 to Layer 2 and to continue to allow LAN administrators to configure and manage their mainframe network topology using the same techniques as their non-mainframe topology.

OSA/SF Virtual MAC and VLAN id Display Capability

The Open Systems Adapter/Support Facility (OSA/SF) has the capability to support virtual Medium Access Control (MAC) and Virtual Local Area Network (VLAN) identifications (IDs) associated with OSA-Express2 feature configured as a Layer 2 interface. This information will now be displayed as a part of an OSA Address Table (OAT) entry. This information is independent of IPv4 and IPv6 formats. There can be multiple Layer 2 VLAN Ids associated to a single unit address. One group MAC can be associated to multiple unit addresses.

For additional information, view IBM Redbooks, IBM System z Connectivity Handbook (SG24-5444) at: www.redbooks.ibm.com/.

HiperSockets

The HiperSockets function, also known as internal Queued Direct Input/Output (iDQIO) or internal QDIO, is an integrated function of the z10 EC server that provides users with attachments to up to sixteen high-speed “virtual” Local Area Networks (LANs) with minimal system and network overhead. HiperSockets eliminates the need to utilize I/O subsystem operations and the need to traverse an external network connection to communicate between logical partitions in the same z10 EC server.

Now, the HiperSockets internal networks on z10 EC can support two transport modes: Layer 2 (Link Layer) as well as the current Layer 3 (Network or IP Layer). Traffic can be Internet Protocol (IP) version 4 or version 6 (IPv4, IPv6) or non-IP (AppleTalk, DECnet, IPX, NetBIOS, or SNA). HiperSockets devices are now protocol-independent and Layer 3 independent. Each HiperSockets device has its own Layer 2 Media Access Control (MAC) address, which is designed to allow the use of applications that depend on the existence of Layer 2 addresses such as DHCP servers and firewalls.

Layer 2 support can help facilitate server consolidation. Complexity can be reduced, network configuration is simplified and intuitive, and LAN administrators can configure and maintain the mainframe environment the same as they do a non-mainframe environment. With support of the new Layer 2 interface by HiperSockets, packet forwarding decisions are now based upon Layer 2 information, instead of Layer 3 information. The HiperSockets device performs automatic MAC address generation and assignment to allow uniqueness within and across logical partitions (LPs) and servers. MAC addresses can also be locally administered. The use of Group MAC addresses for multicast is supported as well as broadcasts to all other Layer 2 devices on the same HiperSockets network. Datagrams are only delivered between HiperSockets devices that are using the same transport mode (Layer 2

with Layer 2 and Layer 3 with Layer 3). A Layer 2 device cannot communicate directly with a Layer 3 device in another LPAR.

A HiperSockets device can filter inbound datagrams by Virtual Local Area Network identification (VLAN ID, IEEE 802.1q), the Ethernet destination MAC address, or both. Filtering can help reduce the amount of inbound traffic being processed by the operating system, helping to reduce CPU utilization.

Analogous to the respective Layer 3 functions, HiperSockets Layer 2 devices can be configured as primary or secondary connectors or multicast routers. This is designed to enable the creation of high performance and high availability Link Layer switches between the internal HiperSockets network and an external Ethernet or to connect the HiperSockets Layer 2 networks of different servers. The HiperSockets Multiple Write Facility for z10 EC is also supported for Layer 2 HiperSockets devices, thus allowing performance improvements for large Layer 2 datastreams.

HiperSockets Layer 2 support is exclusive to System z10 and is supported by z/OS, Linux on System z environments, and z/VM for Linux guest exploitation.

HiperSockets Multiple Write Facility for increased performance

Though HiperSockets provides high-speed internal TCP/IP connectivity between logical partitions within a System z server – the problem is that HiperSockets draws excessive CPU utilization for large outbound messages. This may lead to increased software licensing cost – HiperSockets large outbound messages are charged to a general CPU which can incur high general purpose CPU costs. This may also lead to some performance issues due to synchronous application blocking – HiperSockets large outbound messages will block a sending application while synchronously moving data.

A solution is HiperSockets Multiple Write Facility.

HiperSockets performance has been enhanced to allow for the streaming of bulk data over a HiperSockets link between logical partitions (LPARs). The receiving LPAR can now process a much larger amount of data per I/O interrupt. This enhancement is transparent to the operating system in the receiving LPAR. HiperSockets Multiple Write Facility, with fewer I/O interrupts, is designed to reduce CPU utilization of the sending and receiving LPAR.

The HiperSockets Multiple Write solution moves multiple output data buffers in one write operation.

If the function is disabled then one output data buffer is moved in one write operation. This is also how HiperSockets functioned in the past.

If the function is enabled then multiple output data buffers are moved in one write operation. This reduces CPU utilization related to large outbound messages. When enabled, HiperSockets Multiple Write will be used anytime a message spans an IQD frame requiring multiple output data buffers (SBALs) to transfer the message. Spanning multiple output data buffers can be affected by a number of factors including:

- *IQD frame size*
- *Application socket send size*
- *TCP send size*
- *MTU size*

The HiperSockets Multiple Write Facility is supported in the z/OS environment. For a complete description of the System z10 connectivity capabilities refer to IBM System z Connectivity Handbook, SG24-5444.

HiperSockets Enhancement for zIIP Exploitation

In z/OS V1.10, specifically, the z/OS Communications Server allows the HiperSockets Multiple Write Facility processing for outbound large messages originating from z/OS to be performed on a zIIP. The combination of HiperSockets Multiple Write Facility and zIIP enablement is described as “zIIP-Assisted HiperSockets for large messages.” zIIP-Assisted HiperSockets can help make highly secure, available, virtual HiperSockets networking a more attractive option. z/OS application workloads based on XML, HTTP, SOAP, Java, etc., as well as traditional file transfer, can benefit from zIIP enablement by helping to lower general purpose processor utilization for such TCP/IP traffic.

Only outbound z/OS TCP/IP large messages which originate within a z/OS host are eligible for HiperSockets zIIP-Assisted processing. Other types of network traffic such as IP forwarding, Sysplex Distributor, inbound processing, small messages, or other non TCP/IP network protocols are not eligible for zIIP-Assisted HiperSockets. When the workload is eligible, then the TCP/IP HiperSockets device driver layer (write) processing is redirected to a zIIP, which will unblock the sending application. zIIP-Assisted HiperSockets for large messages is available with z/OS V1.10 with PTF and System z10 only. This feature is unsupported if z/OS is running as a guest in a z/VM environment and is supported for large outbound messages only.

To estimate potential offload, use PROJECTCPU for current and existing workloads. This is accurate and very simple, but you have to be on z/OS 1.10 with the enabling PTFs AND System z10 server AND you need to be performing HiperSockets Multiple Write workload already on z/OS.

Security

Today's world mandates that your systems are secure and available 24/7. The z10 EC employs some of the most advanced security technologies in the industry—helping you to meet rigid regulatory requirements that include encryption solutions, access control management, and extensive auditing features. It also provides disaster recovery configurations and is designed to deliver 99.999% application availability to help avoid the downside of planned downtime, equipment failure, or the complete loss of a data center.

When you need to be more secure, more resilient — **z Can Do IT**. The z10 processor chip has on board cryptographic functions. Standard clear key integrated cryptographic coprocessors provide high speed cryptography for protecting data in storage. CP Assist for Cryptographic Function (CPACF) supports DES, TDES, Secure Hash Algorithms (SHA) for up to 512 bits, Advanced Encryption Standard (AES) for up to 256 bits and Pseudo Random Number Generation (PRNG). Logging has been added to the TKE workstation to enable better problem tracking.

System z is investing in accelerators that provide improved performance for specialized functions. The Crypto Express2 feature for cryptography is an example. The Crypto Express2 feature can be configured as a secure key coprocessor or for Secure Sockets Layer (SSL) acceleration. The feature includes support for 13, 14, 15, 16, 17, 18 and 19 digit Personal Account Numbers for stronger protection of data. And the tamper-resistant cryptographic coprocessor is certified at FIPS 140-2 Level 4.

In 2008, the z10 EC received Common Criteria Evaluation Assurance Level 5 (EAL5) certification for security of logical partitions. System z security is one of the many reasons why the world's top banks and retailers rely on the IBM mainframe to help secure sensitive business transactions.

z Can Do IT securely.

Cryptography

The z10 EC includes both standard cryptographic hardware and optional cryptographic features for flexibility and growth capability. IBM has a long history of providing hardware cryptographic solutions, from the development of Data Encryption Standard (DES) in the 1970s to delivering integrated cryptographic hardware in a server to achieve the US Government's highest FIPS 140-2 Level 4 rating for secure cryptographic hardware.

The IBM System z10 EC cryptographic functions include the full range of cryptographic operations needed for e-business, e-commerce, and financial institution applications. In addition, custom cryptographic functions can be added to the set of functions that the z10 EC offers.

New integrated clear key encryption security features on z10 EC include support for a higher advanced encryption standard and more secure hashing algorithms. Performing these functions in hardware is designed to contribute to improved performance.

Enhancements to eliminate preplanning in the cryptography area include the System z10 function to dynamically add Crypto to a logical partition. Changes to image profiles, to support Crypto Express2 features, are available without an outage to the logical partition. Crypto Express2 features can also be dynamically deleted or moved.

CP Assist for Cryptographic Function (CPACF)

CPACF supports clear-key encryption. All CPACF functions can be invoked by problem state instructions defined by an extension of System z architecture. The function is activated using a no-charge enablement feature and offers the following on every CPACF that is shared between two Processor Units (PUs) and designated as CPs and/or Integrated Facility for Linux (IFL):

- *DES, TDES, AES-128, AES-192, AES-256*
- *SHA-1, SHA-224, SHA-256, SHA-384, SHA-512*
- *Pseudo Random Number Generation (PRNG)*

Enhancements to CP Assist for Cryptographic Function (CPACF):

CPACF has been enhanced to include support of the following on CPs and IFLs:

- *Advanced Encryption Standard (AES) for 192-bit keys and 256-bit keys*
- *SHA-384 and SHA-512 bit for message digest*

SHA-1, SHA-256, and SHA-512 are shipped enabled and do not require the enablement feature.

Support for CPACF is also available using the Integrated Cryptographic Service Facility (ICSF). ICSF is a component of z/OS, and is designed to transparently use the available cryptographic functions, whether CPACF or Crypto Express2, to balance the workload and help address the bandwidth requirements of your applications.

The enhancements to CPACF are exclusive to the System z10 and supported by z/OS, z/VM, z/VSE, and Linux on System z.

Configurable Crypto Express2

The Crypto Express2 feature has two PCI-X adapters. Each of the PCI-X adapters can be defined as either a **Coprocessor** or an **Accelerator**.

Crypto Express2 Coprocessor – for secure-key encrypted transactions (default) is:

- *Designed to support security-rich cryptographic functions, use of secure-encrypted-key values, and User Defined Extensions (UDX)*
- *Designed to support secure and clear-key RSA operations*
- *The tamper-responding hardware and lower-level firmware layers are validated to U.S. Government FIPS 140-2 standard: Security Requirements for Cryptographic Modules at Level 4.*

Crypto Express2 Accelerator – for Secure Sockets Layer (SSL) acceleration:

- *Is designed to support clear-key RSA operations*
- *Offloads compute-intensive RSA public-key and private-key cryptographic operations employed in the SSL protocol. Crypto Express2 features can be carried forward on an upgrade to the System z10 EC, so users may continue to take advantage of the SSL performance and the configuration capability.*

The configurable Crypto Express2 feature is supported by z/OS, z/VM, z/VSE, and Linux on System z. z/VSE offers support for clear-key operations only. Current versions of z/OS, z/VM, and Linux on System z offer support for both clear-key and secure-key operations.

Additional cryptographic functions and features with Crypto Express2

Key management – Added key management for remote loading of ATM and Point of Sale (POS) keys. The elimination of manual key entry is designed to reduce downtime due to key entry errors, service calls, and key management costs.

Improved key exchange – Added Improved key exchange with non-CCA cryptographic systems.

New features added to IBM Common Cryptographic Architecture (CCA) are designed to enhance the ability to exchange keys between CCA systems, and systems that do not use control vectors by allowing the CCA system owner to define permitted types of key import and export while preventing uncontrolled key exchange that can open the system to an increased threat of attack.

These are supported by z/OS and by z/VM for guest exploitation.

Support for ISO 16609

Support for ISO 16609 CBC Mode T-DES Message Authentication (MAC) requirements ISO 16609 CBC Mode T-DES MAC is accessible through ICSF function calls made in the PCI-X Cryptographic Adapter segment 3 Common Cryptographic Architecture (CCA) code.

This is supported by z/OS and by z/VM for guest exploitation.

Support for RSA keys up to 4096 bits

The RSA services in the CCA API are extended to support RSA keys with modulus lengths up to 4096 bits. The services affected include key generation, RSA-based key management, digital signatures, and other functions related to these.

Refer to the ICSF Application Programmers Guide, SA22-7522, for additional details.

Cryptographic enhancements to Crypto Express2

Dynamically add crypto to a logical partition

Today, users can preplan the addition of Crypto Express2 features to a logical partition (LP) by using the Crypto page in the image profile to define the Cryptographic Candidate List, Cryptographic Online List, and Usage and Control Domain Indexes in advance of crypto hardware installation.

With the change to dynamically add crypto to a logical partition, changes to image profiles, to support Crypto Express2 features, are available without outage to the logical partition. Users can also dynamically delete or move Crypto Express2 features. Preplanning is no longer required.

This enhancement is supported by z/OS, z/VM for guest exploitation, z/VSE, and Linux on System z.

Secure Key AES

The Advanced Encryption Standard (AES) is a National Institute of Standards and Technology specification for the encryption of electronic data. It is expected to become the accepted means of encrypting digital information, including financial, telecommunications, and government data.

AES is the symmetric algorithm of choice, instead of Data Encryption Standard (DES) or Triple-DES, for the encryption and decryption of data. The AES encryption algorithm will be supported with secure (encrypted) keys of 128, 192, and 256 bits. The secure key approach, similar to what is supported today for DES and TDES, provides the ability to keep the encryption keys protected at all times, including the ability to import and export AES keys, using RSA public key technology.

Support for AES encryption algorithm includes the master key management functions required to load or generate AES master keys, update those keys, and re-encipher key tokens under a new master key.

Support for 13- thru 19-digit Personal Account Numbers

Credit card companies sometimes perform card security code computations based on Personal Account Number (PAN) data. Currently, ICSF callable services CSNBCSV (VISA CVV Service Verify) and CSNBMSG (VISA CVV Service Generate) are used to verify and to generate a VISA Card Verification Value (CVV) or a MasterCard Card Verification Code (CVC). The ICSF callable services currently support 13-, 16-, and 19-digit PAN data. To provide additional flexibility, new keywords PAN-14, PAN-15, PAN-17, and PAN-18 are implemented in the rule array for both CSNBMSG and CSNBCSV to indicate that the PAN data is comprised of 14, 15, 17, or 18 PAN digits, respectively.

Support for 13- through 19-digit PANs is exclusive to System z10 and is offered by z/OS and z/VM for guest exploitation.

TKE 5.3 workstation and continued support for Smart Card Reader

The Trusted Key Entry (TKE) workstation and the TKE 5.3 level of Licensed Internal Code are optional features on the System z10 EC. The TKE 5.3 Licensed Internal Code (LIC) is loaded on the TKE workstation prior to shipment. The TKE workstation offers security-rich local and remote key management, providing authorized persons a method of operational and master key entry, identification, exchange, separation, and update. The TKE workstation supports connectivity to an Ethernet Local Area Network (LAN) operating at 10 or 100 Mbps. Up to ten TKE workstations can be ordered.

Enhancement with TKE 5.3 LIC

The TKE 5.3 level of LIC includes support for the AES encryption algorithm, adds 256-bit master keys, and includes the master key management functions required to load or generate AES master keys to cryptographic coprocessors in the host.

Also included is an imbedded screen capture utility to permit users to create and to transfer TKE master key entry instructions to diskette or DVD. Under 'Service Management' a "Manage Print Screen Files" utility will be available to all users.

The TKE workstation and TKE 5.3 LIC are available on the z10 EC, z10 BC, z9 EC, and z9 BC.

Smart Card Reader

Support for an optional Smart Card Reader attached to the TKE 5.3 workstation allows for the use of smart cards that contain an embedded microprocessor and associated memory for data storage. Access to and the use of confidential data on the smart cards is protected by a user-defined Personal Identification Number (PIN).

TKE 5.3 LIC has added the capability to store key parts on DVD-RAMs and continues to support the ability to store key parts on paper, or optionally on a smart card. TKE 5.3 LIC has limited the use of floppy diskettes to read-only. The TKE 5.3 LIC can remotely control host cryptographic coprocessors using a password-protected authority signature key pair either in a binary file or on a smart card.

The Smart Card Reader, attached to a TKE workstation with the 5.3 level of LIC will support System z10 BC, z10 EC, z9 EC, and z9 BC. However, TKE workstations with 5.0, 5.1 and 5.2 LIC must be upgraded to TKE 5.3 LIC.

TKE additional smart cards

You have the capability to order Java-based blank smart cards which offers a highly efficient cryptographic and data management application built-in to read-only memory for storage of keys, certificates, passwords, applications, and data. The TKE blank smart cards are compliant with FIPS 140-2 Level 2. When you place an order for a quantity of one, you are shipped 10 smart cards.

System z10 EC cryptographic migration:

Clients using a User Defined Extension (UDX) of the Common Cryptographic Architecture should contact their UDX provider for an application upgrade before ordering a new System z10 EC machine; or before planning to migrate or activate a UDX application to firmware driver level 73 and higher.

- *The Crypto Express2 feature is supported on the System z9 and can be carried forward on an upgrade to the System z10 EC*
- *You may continue to use TKE workstations with 5.3 licensed internal code to control the System z10 EC*
- *TKE 5.0 and 5.1 workstations may be used to control z9 EC, z9 BC, z890, and z990 servers*

Remote Loading of Initial ATM Keys

Typically, a new ATM has none of the financial institution's keys installed. Remote Key Loading refers to the process of loading Data Encryption Standard (DES) keys to Automated Teller Machines (ATMs) from a central administrative site without the need for personnel to visit each machine to manually load DES keys. This has been done by manually loading each of the two clear text key parts individually and separately into ATMs. Manual entry of keys is one of the most error-prone and labor-intensive activities that occur during an installation, making it expensive for the banks and financial institutions.

Remote Key Loading Benefits

- *Provides a mechanism to load initial ATM keys without the need to send technical staff to ATMs*
- *Reduces downtime due to key entry errors*
- *Reduces service call and key management costs*
- *Improves the ability to manage ATM conversions and upgrades*

Integrated Cryptographic Service Facility (ICSF), together with Crypto Express2, support the basic mechanisms in Remote Key Loading. The implementation offers a secure bridge between the highly secure Common Cryptographic Architecture (CCA) environment and the various formats and encryption schemes offered by the ATM vendors. The following ICSF services are offered for Remote Key loading:

- **Trusted Block Create (CSNDTBC)** *This callable service is used to create a trusted block containing a public key and some processing rules.*
- **Remote Key Export (CSNDRKX)** *This callable service uses the trusted block to generate or export DES keys for local use and for distribution to an ATM or other remote device.*

Refer to Application Programmers Guide, SA22-7522, for additional details.

Improved Key Exchange With Non-CCA Cryptographic Systems

IBM Common Cryptographic Architecture (CCA) employs Control Vectors to control usage of cryptographic keys. Non-CCA systems use other mechanisms, or may use keys that have no associated control information. This enhancement provides the ability to exchange keys between CCA systems, and systems that do not use Control Vectors. Additionally, it allows the CCA system owner to define permitted types of key import and export which can help to prevent uncontrolled key exchange that can open the system to an increased threat of attack.

These enhancements are exclusive to System z10, and System z9 and are supported by z/OS and z/VM for z/OS guest exploitation.

On Demand Capabilities

It may sound revolutionary, but it's really quite simple. In the highly unpredictable world of On Demand business, you should get what you need, when you need it. And you should pay for only what you use. Radical? Not to IBM. It's the basic principle underlying IBM capacity on demand for the IBM System z10.

Changes have been made to enhance the Capacity on Demand (CoD) experience for System z10 EC customers:

- *The number of temporary records that can be installed on the Central Processor Complex (CPC) has increased from four to eight.*
- *Resource tokens are now available for On/Off CoD.*

The z10 EC also introduces a architectural approach for temporary offerings that can change the thinking about on demand capacity. One or more flexible configuration definitions can be used to solve multiple temporary situations and multiple capacity configurations can be active at once (for example, activation of just two CBUs out of a definition that has four CBUs is acceptable). This means that On/Off CoD can be active and up to seven other offerings can be active simultaneously. Tokens can be purchased for On/Off CoD so hardware activations can be prepaid.

All activations can be done without having to interact with IBM—when it is determined that capacity is required, no passwords or phone connections are necessary. As long as the total z10 EC can support the maximums that are defined, then they can be made available. With the z10 EC, it is now possible to add permanent capacity while a temporary capacity is currently activated, without having to return first to the original configuration.

Capacity on Demand – Temporary Capacity:

The set of contract documents which support the various Capacity on Demand offerings available for z10 EC has been completely refreshed. While customers with exist-

ing contracts for Capacity Back Up (CBU) and Customer Initiated Upgrade (CIU) – On/Off Capacity on Demand (On/Off CoD) may carry those contracts forward to z10 EC machines, new CoD capability and offerings for z10 EC is only supported by this new contract set.

The new contract set is structured in a modular, hierarchical approach. This new approach will eliminate redundant terms between contract documents, simplifying the contracts for our customers and IBM.

Just-in-time deployment of System z10 EC Capacity on Demand (CoD) is a radical departure from previous System z and zSeries servers. This new architecture allows:

- *Up to eight temporary records to be installed on the CPC and active at any given time*
- *Up to 200 temporary records to be staged on the SE*
- *Variability in the amount of resources that can be activated per record*
- *The ability to control and update records independent of each other*
- *Improved query functions to monitor the state of each record*
- *The ability to add capabilities to individual records concurrently, eliminating the need for constant ordering of new temporary records for different user scenarios*
- *Permanent LIC-CC upgrades to be performed while temporary resources are active*

These capabilities allow you to access and manage processing capacity on a temporary basis, providing increased flexibility for on demand environments. The CoD offerings are built from a common Licensed Internal Code – Configuration Code (LIC-CC) record structure. These Temporary Entitlement Records (TERs) contain the information necessary to control which type of resource can be accessed and to what extent, how many times and for how long, and under what condition – test or real workload. Use of this information gives the different offerings their personality.

Capacity Back Up (CBU): Temporary access to dormant processing units (PUs), intended to replace capacity lost within the enterprise due to a disaster. CP capacity or any and all specialty engine types (zIIP, zAAP, SAP, IFL, ICF) can be added up to what the physical hardware model can contain for up to 10 days for a test activation or 90 days for a true disaster recovery.

On system z10 the CBU entitlement records contain an expiration date that is established at the time of order and is dependent upon the quantity of CBU years. You will now have the capability to extend your CBU entitlements through the purchase of additional CBU years. The number of CBU years per instance of CBU entitlement remains limited to five and fractional years are rounded up to the near whole integer when calculating this limit. For instance, if there are two years and eight months to the expiration date at the time of order, the expiration date can be extended by no more than two additional years. One test activation is provided for each additional CBU year added to the CBU entitlement record.

CBU Tests: The allocation of the default number of test activations changed. Rather than a fixed default number of five test activations for each CBU entitlement record, the number of test activations per instance of the CBU entitlement record will coincide with the number of CBU years, the number of years assigned to the CBU record. This equates to one test activation per year for each CBU entitlement purchased.

These changes apply only to System z10 and to CBU entitlements purchased through the IBM sales channel or directly from Resource Link.

There are now terms governing System z Capacity Back Up (CBU) which allow customers to execute production workload on a CBU Upgrade during a CBU Test..

While all new CBU contract documents contain the new CBU Test terms, existing CBU customers will need to execute a contract to expand their authorization for CBU Test upgrades if they want to have the right to execute production workload on the CBU Upgrade during a CBU Test.

Amendment for CBU Tests

The modification of CBU Test terms is available for existing CBU customers via the IBM Customer Agreement Amendment for IBM System z Capacity Backup Upgrade Tests (in the US this is form number Z125-8145). This amendment can be executed at any time, and separate from any particular order.

Capacity for Planned Event (CPE): Temporary access to dormant PUs, intended to replace capacity lost within the enterprise due to a planned event such as a facility upgrade or system relocation. This offering is available only on the System z10. CPE is similar to CBU in that it is intended to replace lost capacity; however, it differs in its scope and intent. Where CBU addresses disaster recovery scenarios that can take up to three months to remedy, CPE is intended for short-duration events lasting up to three days, maximum. Each CPE record, once activated, gives you access to all dormant PUs on the machine that can be configured in any combination of CP capacity or specialty engine types (zIIP, zAAP, SAP, IFL, ICF).

On/Off Capacity on Demand (On/Off CoD): Temporary access to dormant PUs, intended to augment the existing capacity of a given system. On/Off CoD helps you contain workload spikes that may exceed permanent capacity such that Service Level Agreements cannot be met and business conditions do not justify a permanent upgrade. An On/Off CoD record allows you to temporarily add CP capacity or any and all specialty engine types (zIIP, zAAP, SAP, IFL, ICF) up to the following limits:

- *The quantity of temporary CP capacity ordered is limited by the quantity of purchased CP capacity (permanently active plus unassigned).*

- *The quantity of temporary IFLs ordered is limited by quantity of purchased IFLs (permanently active plus unassigned).*
- *Temporary use of unassigned CP capacity or unassigned IFLs will not incur a hardware charge.*
- *The quantity of permanent zIIPs plus temporary zIIPs can not exceed the quantity of purchased (permanent plus unassigned) CPs plus temporary CPs and the quantity of temporary zIIPs can not exceed the quantity of permanent zIIPs.*
- *The quantity of permanent zAAPs plus temporary zAAPs can not exceed the quantity of purchased (permanent plus unassigned) CPs plus temporary CPs and the quantity of temporary zAAPs can not exceed the quantity of permanent zAAPs.*
- *The quantity of temporary ICFs ordered is limited by the quantity of permanent ICFs as long as the sum of permanent and temporary ICFs is less than or equal to 16.*
- *The quantity of temporary SAPs ordered is limited by the quantity of permanent SAPs as long as the sum of permanent and temporary SAPs is less than or equal to 32.*

Although the System z10 E will allow up to eight temporary records of any type to be installed, only one temporary On/Off CoD record may be active at any given time. An On/Off CoD record may be active while other temporary records are active.

Management of temporary capacity through On/Off CoD is further enhanced through the introduction of resource tokens. For CP capacity, a resource token represents an amount of processing capacity that will result in one MSU of SW cost for one day – an MSU-day. For specialty engines, a resource token represents activation of one engine of that type for one day – an IFL-day, a zIIP-day or a zAAP-day. The different resource tokens are contained in separate pools within the On/Off CoD record. The customer, via the Resource Link ordering process, determines

how many tokens go into each pool. Once On/Off CoD resources are activated, tokens will be decremented from their pools every 24 hours. The amount decremented is based on the highest activation level for that engine type during the previous 24 hours.

Resource tokens are intended to help customers bound the hardware costs associated with using On/Off CoD. The use of resource tokens is optional and they are available on either a prepaid or post-paid basis. When prepaid, the customer is billed for the total amount of resource tokens contained within the On/Off CoD record. When post-paid, the total billing against the On/Off CoD record is limited by the total amount of resource tokens contained within the record. Resource Link will provide the customer an ordering wizard to help determine how many tokens they need to purchase for different activation scenarios. Resource tokens within an On/Off CoD record may also be replenished.

Resource Link offers an ordering wizard to help determine how many tokens you need to purchase for different activation scenarios. Resource tokens within an On/Off CoD record may also be replenished. For more information on the use and ordering of resource tokens, refer to the Capacity on Demand Users Guide, SC28-6871.

Capacity Provisioning

Hardware working with software is critical. The activation of On/Off CoD on z10 EC can be simplified or automated by using z/OS Capacity Provisioning (available with z/OS V1.10 and z/OS V1.9). This capability enables the monitoring of multiple systems based on Capacity Provisioning and Workload Manager (WLM) definitions. When the defined conditions are met, z/OS can suggest capacity changes for manual activation from a z/OS console or the system can add or remove temporary capacity automatically and without operator intervention. **z10 EC can do IT better.**

z/OS Capacity provisioning allows you to set up rules defining the circumstances under which additional capacity should be provisioned in order to fulfill a specific business need. The rules are based on criteria, such as: a specific application, the maximum additional capacity that should be activated, time and workload conditions. This support provides a fast response to capacity changes and ensures sufficient processing power will be available with the least possible delay even if workloads fluctuate.

An installed On/Off CoD record is a necessary prerequisite for automated control of temporary capacity through z/OS Capacity Provisioning.

See z/OS MVS Capacity Provisioning User's Guide (SA33-8299) for more information.

On/Off CoD Test: On/Off CoD allows for a no-charge test. No IBM charges are assessed for the test, including IBM charges associated with temporary hardware capacity, IBM software, or IBM maintenance. This test can be used to validate the processes to download, stage, install, activate, and deactivate On/Off CoD capacity non-disruptively. Each On/Off CoD-enabled server is entitled to only one no-charge test. This test may last up to a maximum duration of 24 hours commencing upon the activation of any capacity resources contained in the On/Off CoD record. Activation levels of capacity may change during the 24 hour test period. The On/Off CoD test automatically terminates at the end of the 24 hours period. In addition to validating the On/Off CoD function within your environment, you may choose to use this test as a training session for your personnel who are authorized to activate On/Off CoD.

SNMP API (Simple Network Management Protocol Application Programming Interface) enhancements have also been made for the new Capacity On Demand features. More information can be found in the System z10 Capacity On Demand User's Guide, SC28-6871.

Capacity on Demand – Permanent Capacity

Customer Initiated Upgrade (CIU) facility: When your business needs additional capacity quickly, Customer Initiated Upgrade (CIU) is designed to deliver it. CIU is designed to allow you to respond to sudden increased capacity requirements by requesting a System z10 EC PU and/or memory upgrade via the Web, using IBM Resource Link, and downloading and applying it to your System z10 EC server using your system's Remote Support connection. Further, with the Express option on CIU, an upgrade may be made available for installation as fast as within a few hours after order submission.

Permanent upgrades: Orders (MESs) of all PU types and memory for System z10 EC servers that can be delivered by Licensed Internal Code, Control Code (LIC-CC) are eligible for CIU delivery. CIU upgrades may be performed up to the maximum available processor and memory resources on the installed server, as configured. While capacity upgrades to the server itself are concurrent, your software may not be able to take advantage of the increased capacity without performing an Initial Programming Load (IPL).

	System z9	System z10
Resources	CP, zIIP, zAAP, IFL, ICF	CP, zIIP, zAAP, IFL, ICF, SAP
Offerings	Requires access to IBM/RETAIN® to activate CBU, On/Off CoD One offering at a time	No password required to IBM/RETAIN to activate CBU, On/Off CoD, CPE Multiple offerings active
Permanent upgrades	Requires de-provisioning of temporary capacity first	Concurrent with temporary offerings
Replenishment	No	Yes w/ CBU & On/Off CoD
CBU Tests	5 tests per record	Up to 15 tests per record
CBU Expiration	No expiration	Specific term length
Capacity Provisioning Manager Support	No	Yes

Reliability, Availability, and Serviceability (RAS)

In today's on demand environment, downtime is not only unwelcome—it's costly. If your applications aren't consistently available, your business suffers. The damage can extend well beyond the financial realm into key areas of customer loyalty, market competitiveness and regulatory compliance. High on the list of critical business requirements today is the need to keep applications up and running in the event of planned or unplanned disruptions to your systems.

While some servers are thought of offering weeks or even months of up time, System z thinks of this in terms of achieving years. The z10 EC continues our commitment to deliver improvements in hardware Reliability, Availability and Serviceability (RAS) with every new System z server. They include microcode driver enhancements, dynamic segment sparing for memory and fixed HSA. The z10 EC is a server that can help keep applications up and running in the event of planned or unplanned disruptions to the system.

The System z10 EC is designed to deliver industry leading reliability, availability and security our customers have come to expect from System z servers. System z10 EC RAS is designed to reduce all sources of outages by reducing unscheduled, scheduled and planned outages. Planned outages are further designed to be reduced with the introduction of concurrent I/O drawer add and eliminating pre-planning requirements. These features are designed to reduce the need for a Power-on-Reset (POR) and help eliminate the need to deactivate/activate/IPL a logical partition.

RAS Design Focus

High Availability (HA) – The attribute of a system designed to provide service during defined periods, at acceptable or agreed upon levels and masks UNPLANNED OUTAGES from end users. It employs fault tolerance, automated failure detection, recovery, bypass reconfiguration, testing, problem and change management.

Continuous Operations (CO) – The attribute of a system designed to continuously operate and mask PLANNED OUTAGES from end users. It employs non-disruptive hardware and software changes, non-disruptive configuration and software coexistence.

Continuous Availability (CA) – The attribute of a system designed to deliver non-disruptive service to the end user 7 days a week, 24 HOURS A DAY (there are no planned or unplanned outages). It includes the ability to recover from a site disaster by switching computing to a second site.

Availability Functions

With the z10 EC, significant steps have been taken in the area of server availability with a focus on reducing pre-planning requirements. Pre-planning requirements are minimized by delivering and reserving 16 GB for HSA so the maximum configuration capabilities can be exploited. And with the introduction of the ability to seamlessly include such events as creation of LPARs, inclusion of logical subsystems, changing logical processor definitions in an LPAR, and the introduction of cryptography into an LPAR. Features that carry forward from previous generation processors include the ability to dynamically enable I/O, and the dynamic swapping of processor types.

Hardware System Area (HSA)

Fixed HSA of 16 GB is provided as standard with the z10 EC. The HSA has been designed to eliminate planning for HSA. Preplanning for HSA expansion for configurations will be eliminated as HCD/IOCP will, via the IOCDs process, always reserve:

- *4 Logical Channel Subsystems (LCSS), pre-defined*
- *60 Logical Partitions (LPARs), pre-defined*
- *Subchannel set 0 with 63.75k devices*
- *Subchannel set 1 with 64K-1 devices*
- *Dynamic I/O Reconfiguration – always enabled by default*
- *Concurrent Patch - always enabled by default*
- *Add/Change the number of logical CP, IFL, ICF, zAAP, zIIP, processors per partition and add SAPs to the configuration*
- *Dynamic LPAR PU assignment optimization CPs, ICFs, IFLs, zAAPs, zIIPs, SAPs*
- *Dynamically Add/Remove Crypto (no LPAR deactivation required)*

Enhanced Book Availability

With proper planning, z10 EC is designed to allow a single book, in a multi-book server, to be non-disruptively removed from the server and re-installed during an upgrade or repair action. To minimize the effect on current workloads and applications, you should ensure that you have sufficient inactive physical resources on the remaining books to complete a book removal.

For customers configuring for maximum availability we recommend to purchasing models with one additional book. To ensure you have the appropriate level of memory, you may want to consider the selection of the Flexible Memory Option features to provide additional resources when completing an Enhanced Book Availability action or when considering plan ahead options for the future. Enhanced Book Availability may also provide benefits should you choose not to configure for maximum availability. In these cases, you should have sufficient inactive resources on the remaining books to contain critical workloads while completing a book replacement. Contact your IBM representative to help you determine and plan the proper configuration to support your workloads when using non-disruptive book maintenance.

Enhanced Book Availability is an extension of the support for Concurrent Book Add (CBA) delivered on z990. CBA makes it possible to concurrently upgrade a server by integrating a second, third, or fourth book into the server without necessarily affecting application processing. The following scenarios prior to the availability of EBA would require a disruptive customer outage. With EBA these upgrade and repair procedures can be performed concurrently without interfering with customer operations.

Concurrent Physical Memory Upgrade

Allows one or more physical memory cards on a single book to be added, or an existing card to be upgraded increasing the amount of physical memory in the system.

Concurrent Physical Memory Replacement

Allows one or more defective memory cards on a single book to be replaced concurrent with the operation of the system.

Concurrent Defective Book Replacement

Allows the concurrent repair of a defective book when that book is operating degraded due to errors such as multiple defective processors.

Enhanced Book Availability is exclusive to z10 EC and z9 EC.

Flexible Memory Option

Flexible memory was first introduced on the z9 EC as part of the design changes and offerings to support enhanced book availability. Flexible memory provides the additional resources to maintain a constant level of memory when replacing a book. On z10 EC, the additional resources required for the flexible memory configurations are provided through the purchase of preplanned memory features along with the purchase of your memory entitlement. In most cases, this implementation provides a lower-cost solution compared to z9 EC. Flexible memory configurations are available on Models E26, E40, E56, and E64 only and range from 32 GB to 1136 GB, model dependent.

Redundant I/O Interconnect

z10 EC with Redundant I/O Interconnect is designed to allow you to replace a book or respond to a book failure and retain connectivity to resources. In the event of a failure or customer initiated action such as the replace-

ment of an HCA2-C fanout card or book, the z10 EC is designed to provide access to your I/O devices through another InfiniBand Multiplexer (IFB-MP) to the affected I/O domains. This is exclusive to System z10 EC and z9 EC.

Enhanced Driver Maintenance

One of the greatest contributors to downtime during planned outages is Licensed Internal Code (LIC) updates. When properly configured, z10 EC is designed to permit select planned LIC updates.

A new query function has been added to validate LIC EDM requirements in advance. Enhanced programmatic internal controls have been added to help eliminate manual analysis by the service team of certain exception conditions.

With the z10 EC, PR/SM code has been enhanced to allow multiple EDM 'From' sync points. Automatic apply of EDM licensed internal change requirements is now limited to EDM and the licensed internal code changes update process.

There are several reliability, availability, and serviceability (RAS) enhancements that have been made to the HMC/SE based on the feedback from the System z9 Enhanced Driver Maintenance field experience.

- *Change to better handle intermittent customer network issues*
- *EDM performance improvements*
- *New EDM user interface features to allow for customer and service personnel to better plan for the EDM*
- *A new option to check all licensed internal code which can be executed in advance of the EDM preload or activate*

Dynamic Oscillator Switchover

The z10 EC has two oscillator cards, a primary and a backup. For most cases, should a failure occur on the primary oscillator card, the backup can detect it, switch over,

and provide the clock signal to the system transparently, with no system outage. Previously, in the event of a failure of the active oscillator, a system outage would occur, the subsequent system Power On Reset (POR) would select the backup, and the system would resume operation. Dynamic Oscillator Switchover is exclusive to System z10 EC and System z9.

Transparent Sparing

The z10 EC offers two PUs reserved as spares per server. In the case of processor failure, these spares are used for transparent sparing. On z10 EC sparing happens on a core granularity rather than chip granularity as on z990 and System z9 (for which “chip” equaled “2 cores”).

Concurrent Maintenance

Concurrent Service for I/O features: All the features that plug into the I/O Cage are able to be added and replaced concurrent with system operation. This virtually eliminates any need to schedule outage for service to upgrade the I/O subsystem on this cage.

Upgrade for Coupling Links: z10 EC has concurrent maintenance for the ISC-3 daughter card. Also, Coupling Links can be added concurrently. This eliminates a need for scheduled downtime in the demanding sysplex environment.

Cryptographic feature: The Crypto Express2 feature plugs in the I/O cage and can be added or replaced concurrently with system operation.

Redundant Cage Controllers: The Power and Service Control Network features redundant Cage Controllers for Logic and Power control. This design enables non-disruptive service to the controllers and virtually eliminates customer scheduled outage.

Auto-Switchover for Support Element (SE): The z10 EC has two Support Elements. In the event of failure on the Primary SE, the switchover to the backup is handled automatically. There is no need for any intervention by the Customer or Service Representative.

Concurrent Memory Upgrade

This function allows adding memory concurrently, up to the maximum amount physically installed. In addition, the Enhanced Book Availability function also enables a memory upgrade to an installed z10 EC book in a multi-book server.

Plan Ahead Memory

Future memory upgrades can now be preplanned to be non-disruptive. The preplanned memory feature will add the necessary physical memory required to support target memory sizes. The granularity of physical memory in the System z10 design is more closely associated with the granularity of logical, entitled memory, leaving little room for growth. If you anticipate an increase in memory requirements, a “target” logical memory size can now be specified in the configuration tool along with a “starting” logical memory size. The configuration tool will then calculate the physical memory required to satisfy this target memory. Should additional physical memory be required, it will be fulfilled with the currently available preplanned memory features.

The preplanned memory feature is offered in 16 gigabyte (GB) increments. The quantity assigned by the configuration tool is the number of 16 GB blocks necessary to increase the physical memory from that required for the “starting” logical memory to the physical memory required for the “target” logical configuration. Activation of any preplanned memory requires the purchase of a preplanned

Environmental Enhancements

memory activation feature. One pre-planned memory activation feature is required for each preplanned memory feature. You now have the flexibility to activate memory to any logical size offered between the starting and target size.

Plan ahead memory is exclusive to System z10 and is transparent to operating systems.

Service Enhancements

z10 EC service enhancements designed to avoid scheduled outages include:

- *Concurrent firmware fixes*
- *Concurrent driver upgrades*
- *Concurrent parts replacement*
- *Concurrent hardware upgrades*
- *DIMM FRU indicators*
- *Single processor core checkstop*
- *Single processor core sparing*
- *Point-to-Point SMP Fabric (not a ring)*
- *FCP end-to-end checking*
- *Hot swap of ICB-4 and InfiniBand hub cards*
- *Redundant 100 Mb Ethernet service network with VLAN*

Power and cooling discussions have entered the budget planning of every IT environment. As energy prices have risen and utilities have restricted the amount of power usage, it is important to review the role of the server in balancing IT spending.

Power Monitoring

The “mainframe gas gauge” feature introduced on the System z9 servers, provides power and thermal information via the System Activity Display (SAD) on the Hardware Management Console and will be available on the z10 EC giving a point in time reference of the information. The current total power consumption in watts and BTU/hour as well as the air input temperature will be displayed.

Power Estimation Tool

To assist in energy planning, Resource Link provides tools to estimate server energy requirements before a new server purchase. A user will input the machine model, memory, and I/O configuration and the tool will output an estimate of the system total heat load and utility input power. A customized planning aid is also available on Resource Link which provides physical characteristics of the machine along with cooling recommendations, environmental specifications, system power rating, power plugs/receptacles, line cord wire specifications and the machine configuration.

Parallel Sysplex Cluster Technology

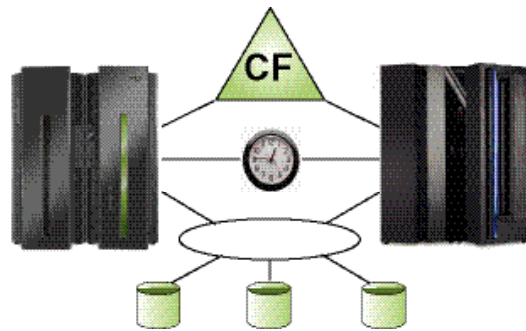
IBM Systems Director Active Energy Manager

IBM Systems Director Active Energy Manager™ (AEM) is a building block which enables customers to manage actual power consumption and resulting thermal loads IBM servers place in the data center. The z10 EC provides support for IBM Systems Director Active Energy Manager (AEM) for Linux on System z for a single view of actual energy usage across multiple heterogeneous IBM platforms within the infrastructure. AEM for Linux on System z will allow tracking of trends for both the z10 EC as well as multiple server platforms. With this trend analysis, a data center administrator will have the data to help properly estimate power inputs and more accurately plan data center consolidation or modification projects.

On System z10, the HMC will now provide support for the Active Energy Manager (AEM) which will display power consumption/air input temperature as well as exhaust temperature. AEM will also provide some limited status configuration information which might assist in explaining changes to the power consumption. AEM is exclusive to System z10.

IBM System z servers stand alone against competition and have stood the test of time with our business resiliency solutions. Our coupling solutions with Parallel Sysplex technology allow for greater scalability and availability.

Parallel Sysplex clustering is designed to bring the power of parallel processing to business-critical System z10, System z9, z990 or z890 applications. A Parallel Sysplex cluster consists of up to 32 z/OS images coupled to one or more Coupling Facilities (CFs or ICFs) using high-speed specialized links for communication. The Coupling Facilities, at the heart of the Parallel Sysplex cluster, enable high speed, read/write data sharing and resource sharing among all the z/OS images in a cluster. All images are also connected to a Sysplex Timer® or by implementing the Server Time Protocol (STP), so that all events can be properly sequenced in time.



Parallel Sysplex Resource Sharing enables multiple system resources to be managed as a single logical resource shared among all of the images. Some examples of resource sharing include JES2 Checkpoint, GRS “star,” and Enhanced Catalog Sharing; all of which provide simplified systems management, increased performance and/or scalability.

Although there is significant value in a single footprint and multi-footprint environment with resource sharing, those customers looking for high availability must move on to a database data sharing configuration. With the Parallel Sysplex environment, combined with the Workload Manager and CICS TS, DB2 or IMS, incoming work can

be dynamically routed to the z/OS image most capable of handling the work. This dynamic workload balancing, along with the capability to have read/write access data from anywhere in the Parallel Sysplex cluster, provides scalability and availability. When configured properly, a Parallel Sysplex cluster is designed with no single point of failure and can provide customers with near continuous application availability over planned and unplanned outages.

With the introduction of the z10 EC, we have the concept of n-2 on the hardware as well as the software. The z10 EC participates in a Sysplex with System z10 BC, System z9, z990 and z890 only and currently supports z/OS 1.8 and higher.

For detailed information on IBM's Parallel Sysplex technology, visit our Parallel Sysplex home page at <http://www-03.ibm.com/systems/z/psol/>.

Coupling Facility Control Code (CFCC) Level 16

CFCC Level 16 is being made available on the IBM System z10 EC.

Improved service time with Coupling Facility Duplexing enhancements:

Prior to Coupling Facility Control Code (CFCC) Level 16, System-Managed Coupling Facility (CF) Structure Duplexing required two duplexing protocol exchanges to occur synchronously during processing of each duplexed structure request. CFCC Level 16 allows one of these protocol exchanges to complete asynchronously. This allows faster duplexed request service time, with more benefits when the Coupling Facilities are further apart, such as in a multi-site Parallel Sysplex environment.

List notification improvements: Prior to CFCC Level 16, when a shared queue (subsidiary list) changed state from empty to non-empty, the CF would notify ALL active connectors. The first one to respond would process the new message, but when the others tried to do the same, they would find nothing, incurring additional overhead.

CFCC Level 16 can help improve the efficiency of coupling communications for IMS Shared Queue and WebSphere MQ Shared Queue environments. The Coupling Facility notifies only one connector in a sequential fashion. If the shared queue is processed within a fixed period of time, the other connectors do not need to be notified, saving the cost of the false scheduling. If a shared queue is not read within the time limit, then the other connectors are notified as they were prior to CFCC Level 16.

When migrating CF levels, lock, list and cache structure sizes might need to be increased to support new function. For example, when you upgrade from CFCC Level 15 to Level 16 the required size of the structure might increase. This adjustment can have an impact when the system allocates structures or copies structures from one coupling facility to another at different CF levels.

The coupling facility structure sizer tool can size structures for you and takes into account the amount of space needed for the current CFCC levels.

Access the tool at: <http://www.ibm.com/servers/eserver/zseries/cfsizer/>.

CFCC Level 16 is exclusive to System z10 and is supported by z/OS and z/VM for guest exploitation.

Coupling Facility Configuration Alternatives

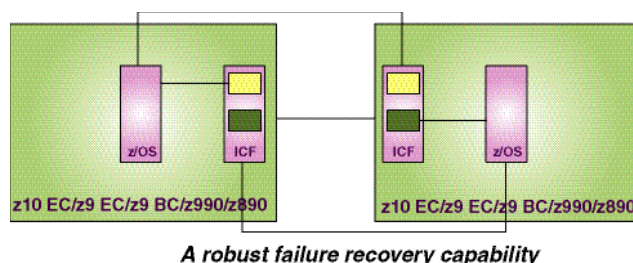
IBM offers multiple options for configuring a functioning Coupling Facility:

- Standalone Coupling Facility:** The standalone CF provides the most “robust” CF capability, as the CPC is wholly dedicated to running the CFCC microcode — all of the processors, links and memory are for CF use only. A natural benefit of this characteristic is that the standalone CF is always failure-isolated from exploiting z/OS software and the server that z/OS is running on for environments without System-Managed CF Structure Duplexing. While there is no unique standalone coupling facility model offered with the z10 EC, customers can achieve the same physically isolated environment as on prior mainframe families by ordering a z10 EC, z9 EC, z9 BC, and z990 with PUs characterized as Internal Coupling Facilities (ICFs). There are no software charges associated with such a configuration.
- Internal Coupling Facility (ICF):** Customers considering clustering technology can get started with Parallel Sysplex technology at a lower cost by using an ICF instead of purchasing a standalone Coupling Facility. An ICF feature is a processor that can only run Coupling Facility Control Code (CFCC) in a partition. Since CF LPARs on ICFs are restricted to running only CFCC, there are no IBM software charges associated with ICFs. ICFs are ideal for Intelligent Resource Director and resource sharing environments as well as for data sharing environments where System-Managed CF Structure Duplexing is exploited.

System-Managed CF Structure Duplexing

System-Managed Coupling Facility (CF) Structure Duplexing provides a general purpose, hardware-assisted, easy-to-exploit mechanism for duplexing CF structure data. This provides a robust recovery mechanism for failures such as loss of a single structure or CF or loss of connectivity to a single CF, through rapid failover to the backup instance of the duplexed structure pair. CFCC Level 16 provides CF Duplex-

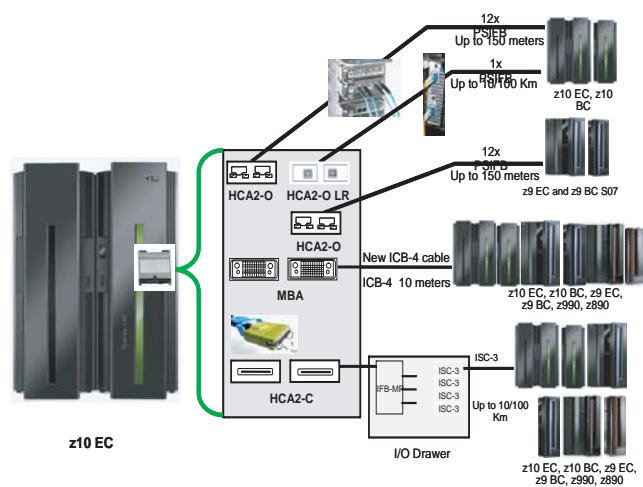
ing enhancements described previously in the section titled “Coupling Facility Control Code (CFCC) Level 16”.



Parallel Sysplex Coupling Connectivity

The Coupling Facilities communicate with z/OS images in the Parallel Sysplex environment over specialized high-speed links. As processor performance increases, it is important to also use faster links so that link performance does not become constrained. The performance, availability and distance requirements of a Parallel Sysplex environment are the key factors that will identify the appropriate connectivity option for a given configuration.

When connecting between System z10, System z9 and z990/z890 servers the links must be configured to operate in Peer Mode. This allows for higher data transfer rates to and from the Coupling Facilities. The peer link acts simultaneously as both a CF Sender and CF Receiver link, reducing the number of links required. Larger and more data buffers and improved protocols may also improve long distance performance.



Introducing long reach InfiniBand coupling links

Now, InfiniBand can be used for Parallel Sysplex coupling and STP communication at unrepeated distances up to 10 km (6.2 miles) and greater distances when attached to qualified optical networking solutions. InfiniBand coupling links supporting extended distance is referred to as Long Reach 1x (one pair of fiber) InfiniBand.

- *Long Reach 1x InfiniBand coupling links support single data rate (SDR) at 2.5 gigabits per second (Gbps) when connected to a DWDM capable of SDR (1x IB-SDR).*
- *Long Reach 1x InfiniBand coupling links support double data rate (DDR) at 5 Gbps when connected to a DWDM capable of DDR (1x IB-DDR).*

The link data rate will auto-negotiate from SDR to DDR depending upon the capability of the attached equipment.

Other advantages of Parallel Sysplex using InfiniBand (PSIFB):

- *InfiniBand coupling links also provide the ability to define up to 16 CHPIDs on a single PSIFB port, allowing physical coupling links to be shared by multiple sysplexes. This also provides additional subchannels for Coupling Facility communication, improving scalability, and reducing contention in heavily utilized system configurations. It also allows for one CHPID to be directed to one CF, and another CHPID directed to another CF on the same target server, using the same port.*
- *Like other coupling links, external InfiniBand coupling links are also valid to pass time synchronization signals for Server Time Protocol (STP). Therefore the same coupling links can be used to exchange timekeeping information and Coupling Facility messages in a Parallel Sysplex environment.*
- *The IBM System z10 EC also takes advantage of InfiniBand as a higher-bandwidth replacement for the Self-Timed Interconnect (STI) I/O interface features found in prior System z servers.*

The IBM System z10 EC will support up to 32 PSIFB links as compared to 16 PSIFB links on System z9 servers. For either z10 EC or z9, there must be less than or equal to a total of 32 PSIFBs and ICB-4 links.

InfiniBand coupling links are CHPID type CIB.

Coupling Connectivity for Parallel Sysplex

You now have five coupling link options for communication in a Parallel Sysplex environment:

1. **Internal Coupling Channels (ICs)** can be used for internal communication between Coupling Facilities (CFs) defined in LPARs and z/OS images on the same server.
2. **Integrated Cluster Bus-4 (ICB-4)** is for short distances. ICB-4 links use 10 meter (33 feet) copper cables, of which 3 meters (10 feet) is used for internal routing and strain relief. ICB-4 is used to connect z10 EC-to-z10 EC, z10 BC, z9 EC, z9 BC, z990, and z890. Note. If connecting to a z10 BC or a z9 BC with ICB-4, those servers cannot be installed with the nonraised floor feature. Also, if the z10 BC is ordered with the nonraised floor feature, ICB-4 cannot be ordered.
3. **12x InfiniBand coupling links (12x IB-SDR or 12x IB-DDR)** offer an alternative to ISC-3 in the data center and facilitate coupling link consolidation. Physical links can be shared by multiple operating system images or Coupling Facility images on a single system. The 12x InfiniBand links support distances up to 150 meters (492 feet) using industry-standard OM3 50 micron multimode fiber optic cables.
4. **Long Reach 1x InfiniBand coupling links (1x IB-SDR or 1x IB-DDR)** are an alternative to ISC-3 and offer greater distances with support for point-to-point unrepeated distances up to 10 km (6.2 miles) using 9 micron single mode fiber optic cables. Greater distances can be supported with System z-qualified optical networking solutions. Long Reach 1x InfiniBand coupling links support the same sharing capabilities as the 12x InfiniBand version, allowing one physical link to be shared by multiple operating system images or Coupling Facility images on a single system.

System z now supports 12x InfiniBand single data rate (12x IB-SDR) coupling link attachment between System z10 and System z9 general purpose (no longer limited to standalone coupling facility)

5. **InterSystem Channel-3 (ISC-3)** supports communication at unrepeated distances up to 10 km (6.2 miles) using 9 micron single mode fiber optic cables and greater distances with System z-qualified optical networking solutions. ISC-3s are supported exclusively in peer mode (CHPID type CFP).

Note: The InfiniBand link data rates do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. Specifically, with 12x InfiniBand coupling links, while the link data rate is higher than that of ICB, the service times of coupling operations are greater, and the actual throughput is less.

Refer to the Coupling Facility Configuration Options whitepaper for a more specific explanation of when to continue using the current ICB or ISC-3 technology versus migrating to InfiniBand coupling links.

The whitepaper is available at: <http://www.ibm.com/systems/z/advantages/pso/whitepaper.html>.

z10 Coupling Link Options

Type	Description	Use	Link data rate	Distance	z10 BC z10 EC Max	z10 Max
PSIFB	1x IB-DDR LR	z10 to z10	5 Gbps	10 km unrepeated (6.2 miles) 100 km repeated	12*/32*	64 CHPIDS
PSIFB	12x IB-DDR	z10 to z10 z10 to z9	6 GBps 3 GBps**	150 meters (492 ft)***	12*/32*	
IC	Internal Coupling Channel	Internal Communication	Internal Speeds	N/A	32/32	
ICB-4	Copper connection between OS and CF	z10, z9 z990, z890	2 GBps	10 meters*** (33 ft)	12/16	
ISC-3	Fiber connection between OS and CF	z10, z9 z990, z890	2 Gbps	10 km unrepeated (6.2 miles) 100 km repeated	48/48	

- The maximum number of Coupling Links combined cannot exceed 64 per server (PSIFB, ICB-4, ISC-3). There is a maximum of 64 Coupling CHPIDs (CIB, ICP, CBP, CFP) per server.

- For each MBA fanout installed for ICB-4s, the number of possible customer HCA fanouts is reduced by one

* Each link supports definition of multiple CIB CHPIDs, up to 16 per fanout

** z10 negotiates to 3 GBps (12x IB-SDR) when connected to a System z9*

** 3 meters (10 feet) reserved for internal routing and strain relief

Note: The InfiniBand link data rates of 6 GBps, 3 GBps, 2.5 Gbps, or 5 Gbps do not represent the performance of the link. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload. With InfiniBand coupling links, while the link data rate may be higher than that of ICB (12x IB-SDR or 12x IB-DDR) or ISC-3 (1x IB-SDR or 1x IB-DDR), the service times of coupling operations are greater, and the actual throughput may be less than with ICB links or ISC-3 links.

Time synchronization and time accuracy on z10 EC

If you require time synchronization across multiple servers (for example you have a Parallel Sysplex environment) or you require time accuracy either for one or more System z servers or you require the same time across heterogeneous platforms (System z, UNIX, AIX®, etc.) you can meet these requirements by either installing a Sysplex Timer Model 2 (9037-002) or by implementing Server Time Protocol (STP).

The Sysplex Timer Model 2 is the centralized time source that sets the Time-Of-Day (TOD) clocks in all attached servers to maintain synchronization. The Sysplex Timer Model 2 provides the stepping signal that helps ensure that all TOD clocks in a multi-server environment increment in unison to permit full read or write data sharing with integrity. The Sysplex Timer Model 2 is a key component of an IBM Parallel Sysplex environment and a GDPS® availability solution for On Demand Business.

The z10 EC server requires the External Time Reference (ETR) feature to attach to a Sysplex Timer. The ETR feature is standard on the z10 EC and supports attachment at an unrepeated distance of up to three kilometers (1.86 miles) and a link data rate of 8 Megabits per second. The distance from the Sysplex Timer to the server can be extended to 100 km using qualified Dense Wavelength Division Multiplexers (DWDMs). However, the maximum repeated distance between Sysplex Timers is limited to 40 km.

Server Time Protocol (STP)

STP is a message-based protocol in which timekeeping information is transmitted between servers over externally defined coupling links. ICB-4, ISC-3, and InfiniBand coupling links can be used to transport STP messages.

Server Time Protocol (STP) Enhancements

STP configuration and time information restoration after Power on Resets (POR) or power outage:

This enhancement delivers system management improvements by restoring the STP configuration and time information after Power on Resets (PORs) or power failure that affects both servers of a two server STP-only Coordinated Timing Network (CTN). To enable this function the customer has to select an option that will assure that no other servers can join the two server CTN. Previously, if both the Preferred Time Server (PTS) and the Backup Time Server (BTS) experienced a simultaneous power outage (site failure), or both experienced a POR, reinitialization of time, and special roles (PTS, BTS, and CTS) was required. With this enhancement, you will no longer need to reinitialize the time or reassign the roles for these events.

Preview - Improved STP System Management with

new z/OS Messaging: This is a new function planned to generate z/OS messages when various hardware events that affect the External Time Sources (ETS) configured for an STP-only CTN occur. This may improve problem determination and correction times. Previously, the messages were generated only on the Hardware Management Console (HMC).

The ability to generate z/OS messages will be supported on IBM System z10 and System z9 servers with z/OS 1.11 (with enabling support rolled back to z/OS 1.9) in the second half of 2009.

The following STP enhancements are available on System z10 and System z9 servers.

The STP feature and the latest Machine Change Levels are required.

Enhanced Network Time Protocol (NTP) client support:

This enhancement addresses the requirements for those who need to provide the same accurate time across heterogeneous platforms in an enterprise.

The STP design has been enhanced to include support for a Simple Network Time Protocol (SNTP) client on the Support Element. By configuring an NTP server as the STP External Time Source (ETS), the time of an STP-only Coordinated Timing Network (CTN) can track to the time provided by the NTP server, and maintain a time accuracy of 100 milliseconds.

Note: NTP client support has been available since October 2007.

Enhanced accuracy to an External Time Source: The time accuracy of an STP-only CTN has been improved by adding the capability to configure an NTP server that has a pulse per second (PPS) output signal as the ETS device. This type of ETS device is available worldwide from several vendors that provide network timing solutions.

STP has been designed to track to the highly stable, accurate PPS signal from the NTP server, and maintain an accuracy of 10 microseconds as measured at the PPS input of the System z server. A number of variables such as accuracy of the NTP server to its time source (GPS, radio signals for example), and cable used to connect the PPS signal will determine the ultimate accuracy of STP to Coordinated Universal Time (UTC).

In comparison, the IBM Sysplex Timer is designed to maintain an accuracy of 100 microseconds when attached to an ETS with a PPS output. If STP is configured to use a dial-out time service or an NTP server without PPS, it is designed to provide a time accuracy of 100 milliseconds to the ETS device.

For this enhancement, the NTP output of the NTP server has to be connected to the Support Element (SE) LAN, and the PPS output of the same NTP server has to be connected to the PPS input provided on the External Time Reference (ETR) feature of the System z10 or System z9 server.

Continuous availability of NTP servers used as External Time Source: Improved External Time Source (ETS) availability can now be provided if you configure different NTP servers for the Preferred Time Server (PTS) and the Backup Time Server (BTS). Only the PTS or the BTS can be the Current Time Server (CTS) in an STP-only CTN. Prior to this enhancement, only the CTS calculated the time adjustments necessary to maintain time accuracy. With this enhancement, if the PTS/CTS cannot access the NTP server or the pulse per second (PPS) signal from the NTP server, the BTS, if configured to a different NTP server, may be able to calculate the adjustment required and propagate it to the PTS/CTS. The PTS/CTS in turn will perform the necessary time adjustment steering.

This avoids a manual reconfiguration of the BTS to be the CTS, if the PTS/CTS is not able to access its ETS. In an ETR network when the primary Sysplex Timer is not able to access the ETS device, the secondary Sysplex Timer takes over the role of the primary - a recovery action not

always accepted by some environments. The STP design provides continuous availability of ETS while maintaining the special roles of PTS and BTS assigned by the enterprise.

The improvement is available when the ETS is configured as an NTP server or an NTP server using PPS.

NTP server on Hardware Management Console (HMC):

Improved security can be obtained by providing NTP server support on the HMC. If an NTP server (with or without PPS) is configured as the ETS device for STP, it needs to be attached directly to the Support Element (SE) LAN. The SE LAN is considered by many users to be a private dedicated LAN to be kept as isolated as possible from the intranet or Internet.

Since the HMC is normally attached to the SE LAN, providing an NTP server capability on the HMC addresses the potential security concerns most users may have for attaching NTP servers to the SE LAN. The HMC, using a separate LAN connection, can access an NTP server available either on the intranet or Internet for its time source. Note that when using the HMC as the NTP server, there is no pulse per second capability available. Therefore, you should not configure the ETS to be an NTP server using PPS.

Enhanced STP recovery when Internal Battery Feature

is in use: Improved availability can be obtained when power has failed for a single server (PTS/CTS), or when there is a site power outage in a multisite configuration where the PTS/CTS is installed (the site with the BTS is a different site not affected by the power outage). If an Internal Battery Feature (IBF) is installed on your System

z server, STP now has the capability of receiving notification that customer power has failed and that the IBF is engaged. When STP receives this notification from a server that has the role of the PTS/CTS, STP can automatically reassign the role of the CTS to the BTS, thus automating the recovery action and improving availability.

STP configuration and time information saved across Power-on-Resets (POR) or power outages:

This enhancement delivers system management improvements by saving the STP configuration across PORs and power failures for a single server STP-only CTN. Previously, if there was a POR of the server or the server experienced a power outage, the time and assignment of the PTS and CTS roles would have to be reinitialized. You will no longer need to reinitialize the time or reassign the role of PTS/CTS across POR or power outage events.

Note: This enhancement is also available on the z990 and z890 servers, in addition to System z10 and System z9 servers.

Application Programming Interface (API) to automate STP CTN reconfiguration:

The concept of “a pair and a spare” has been around since the original Sysplex Couple Data Sets (CDSs). If the primary CDS becomes unavailable, the backup CDS would take over. Many sites have had automation routines bring a new backup CDS online to avoid a single point of failure. This idea is being extended to STP. With this enhancement, if the PTS fails and the BTS takes over as CTS, an API is now available on the HMC so you can automate the reassignment of the PTS, BTS, and Arbiter roles. This can improve availability by avoiding a single point of failure after the BTS has taken over as the CTS.

Prior to this enhancement, the PTS, BTS, and Arbiter roles had to be reassigned manually using the System (Sysplex) Time task on the HMC. For additional details on the API, please refer to System z Application Programming Interfaces, SB10-7030-11.

Additional information is available on the STP Web page: <http://www.ibm.com/systems/z/psso/stp.html>.

The following Redbooks are available at the Redbooks Web site: <http://www.redbooks.ibm.com/>.

- *Server Time Protocol Planning Guide, SG24-7280*
- *Server Time Protocol Implementation Guide, SG24-7281*

Internal Battery Feature Recommendation

Single data center

- *CTN with 2 servers, install IBF on at least the PTS/CTS*
 - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*
 - *CTN with 3 or more servers IBF not required for STP recovery, if Arbiter configured*

Two data centers

- *CTN with 2 servers (one in each data center) install IBF on at least the PTS/CTS*
 - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*
- *CTN with 3 or more servers, install IBF on at least the PTS/CTS*
 - *Also recommend IBF on BTS to provide recovery protection when BTS is the CTS*

Message Time Ordering (Sysplex Timer Connectivity to Coupling Facilities)

As processor and Coupling Facility link technologies have improved, the requirement for time synchronization tolerance between systems in a Parallel Sysplex environment has become ever more rigorous. In order to enable any exchange of time stamped information between systems in a sysplex involving the Coupling Facility to observe the correct time ordering, time stamps are now included in the message-transfer protocol between the systems and the Coupling Facility. Therefore, when a Coupling Facility is configured on any System z10 or System z9, the Coupling Facility will require connectivity to the same 9037 Sysplex Timer or Server Time Protocol (STP) configured Coordinated Timing Network (CTN) that the systems in its Parallel Sysplex cluster are using for time synchronization. If the ICF is on the same server as a member of its Parallel Sysplex environment, no additional connectivity is required, since the server already has connectivity to the Sysplex Timer.

However, when an ICF is configured on any z10 which does not host any systems in the same Parallel Sysplex cluster, it is necessary to attach the server to the 9037 Sysplex Timer or implement STP.

HMC System Support

The new functions available on the Hardware Management Console (HMC) version 2.10.1 apply exclusively to System z10. However, the HMC version 2.10.1 will continue to support System z9, zSeries, and S/390® G5/G6 servers.

The 2.10.1 HMC will continue to support up to two 10 Mbps or 100 Mbps Ethernet LANs. A Token Ring LAN is not supported. The 2.10.1 HMC applications have been updated to support HMC hardware without a diskette drive. DVD-RAM, CD-ROM, and/or USB flash memory drive media will be used.

Family	Machine Type	Firmware Driver	SE Version
z10 BC	2098	76	2.10.1
z10 EC	2097	73	2.10.0
z9 BC	2096	67	2.9.2
z9 EC	2094	67	2.9.2
z890	2086	55	1.8.2
z990	2084	55	1.8.2
z800	2066	3G	1.7.3
z900	2064	3G	1.7.3
9672 G6	9672/9674	26	1.6.2
9672 G5	9672/9674	26	1.6.2

Internet Protocol, Version 6 (IPv6)

HMC version 2.10.1 and Support Element (SE) version 2.10.1 can now communicate using IP Version 4 (IPv4), IP Version 6 (IPv6), or both. It is no longer necessary to assign a static IP address to an SE if it only needs to communicate with HMCs on the same subnet. An HMC and SE can use IPv6 link-local addresses to communicate with each other.

HMC/SE support is addressing the following requirements:

- *The availability of addresses in the IPv4 address space is becoming increasingly scarce*
- *The demand for IPv6 support is high in Asia/Pacific countries since many companies are deploying IPv6*
- *The U.S. Department of Defense and other U.S. government agencies are requiring IPv6 support for any products purchased after June 2008*

More information on the U.S. government requirements can be found at: <http://www.whitehouse.gov/omb/memoranda/fy2005/m05-22.pdf> and http://www.whitehouse.gov/omb/egov/documents/IPv6_FAQs.pdf

HMC/SE Console Messenger

On servers prior to System z9, the remote browser capability was limited to Platform Independent Remote Console (PIRC), with a very small subset of functionality. Full functionality using Desktop-On-Call (DTC) was limited to one user at a time and was slow, so it was rarely used.

With System z9, full functionality to multiple users was delivered with a fast Web browser solution. You liked this, but requested the ability to communicate to other remote users.

There is now a new console messenger task that offers basic messaging capabilities to allow system operators or administrators to coordinate their activities. The new task may be invoked directly, or using a new option in Users and Tasks. This capability is available for HMC and SE

local and remote users permitting interactive plain-text communication between two users and also allowing a user to broadcast a plain-text message to all users. This feature is a limited messenger application and does not interact with other messengers.

HMC z/VM Tower systems management enhancements

Building upon the previous VM systems management support from the Hardware Management Console (HMC), which offered management support for already defined virtual resources, new HMC capabilities are being made available allowing selected virtual resources to be defined. In addition, further enhancements have been made for managing defined virtual resources.

Enhancements are designed to deliver out-of-the-box integrated graphical user interface-based (GUI-based) management of selected parts of z/VM. This is especially targeted to deliver ease-of-use for enterprises new to System z.

This helps to avoid the purchase and installation of additional hardware or software, which may include complicated setup procedures. You can more seamlessly perform hardware and selected operating system management using the HMC Web browser-based user interface.

Enhanced installation support for z/VM using the HMC:

HMC version 2.10.1, along with Support Element (SE) version 2.10.1 on z10 EC, now gives you the ability to install Linux on System z in a z/VM virtual machine using the

HMC DVD drive. This new function does not require an external network connection between z/VM and the HMC, but instead uses the existing communication path between the HMC and the SE.

This support is intended for environments that have no alternative, such as a LAN-based server, for serving the DVD contents for Linux installations. The elapsed time for installation using the HMC DVD drive can be an order of magnitude, or more, longer than the elapsed time for LAN-based alternatives.

Using the current support and the z/VM support, z/VM can be installed in an LPAR and both z/VM and Linux on System z can be installed in a virtual machine from the HMC DVD drive without requiring an external network setup or a connection between an LPAR and the HMC.

This addresses security concerns and additional configuration efforts using the only other previous solution of the external network connection from the HMC to the z/VM image.

Enhanced installation support using the HMC is exclusive to System z10 and is supported by z/VM.

Implementation Services for Parallel Sysplex

IBM Implementation Services for Parallel Sysplex CICS and WAS Enablement

IBM Implementation Services for Parallel Sysplex Middleware – CICS enablement consists of five fixed-price and fixed-scope selectable modules:

- 1) *CICS application review*
- 2) *z/OS CICS infrastructure review (module 1 is a prerequisite for this module)*
- 3) *CICS implementation (module 2 is a prerequisite for this module)*
- 4) *CICS application migration*
- 5) *CICS health check*

IBM Implementation Services for Parallel Sysplex Middleware – WebSphere Application Server enablement consists of three fixed-price and fixed-scope selectable modules:

- 1) *WebSphere Application Server network deployment planning and design*
- 2) *WebSphere Application Server network deployment implementation (module 1 is a prerequisite for this module)*
- 3) *WebSphere Application Server health check*

For a detailed description of this service, refer to Services Announcement 608-041, (RFA47367) dated June 24, 2008.

Implementation Services for Parallel Sysplex DB2 Data Sharing

To assist with the assessment, planning, implementation, testing, and backup and recovery of a System z DB2 data sharing environment, IBM Global Technology Services announced and made available the IBM Implementation Services for Parallel Sysplex Middleware – DB2 data sharing on February 26, 2008.

This DB2 data sharing service is designed for clients who want to:

- 1) *Enhance the availability of data*
- 2) *Enable applications to take full utilization of all servers' resources*
- 3) *Share application system resources to meet business goals*
- 4) *Manage multiple systems as a single system from a single point of control*
- 5) *Respond to unpredicted growth by quickly adding computing power to match business requirements without disruption*
- 6) *Build on the current investments in hardware, software, applications, and skills while potentially reducing computing costs*

The offering consists of six selectable modules; each is a stand-alone module that can be individually acquired. The first module is an infrastructure assessment module, followed by five modules which address the following DB2 data sharing disciplines:

- 1) *DB2 data sharing planning*
- 2) *DB2 data sharing implementation*
- 3) *Adding additional data sharing members*
- 4) *DB2 data sharing testing*
- 5) *DB2 data sharing backup and recovery*

For more information on these services contact your IBM representative or refer to: www.ibm.com/services/server.

GDPS

Geographically Dispersed Parallel Sysplex™ (GDPS) is designed to provide a comprehensive end-to-end continuous availability and/or disaster recovery solution for

Fiber Quick Connect for FICON LX Environments

System z servers. Now Geographically Dispersed Open Clusters (GDOC) is designed to address this need for open systems. GDPS 3.5 will support GDOC for coordinated disaster recovery across System z and non-System z servers if Veritas Cluster Server is already installed. GDPS and the Basic HyperSwap (available with z/OS V1.9) solutions help to ensure system failures are invisible to employees, partners and customers with dynamic disk-swapping capabilities that ensure applications and data are available.

GDPS is a multi-site or single-site end-to-end application availability solution that provides the capability to manage remote copy configuration and storage subsystems (including IBM TotalStorage), to automate Parallel Sysplex operation tasks and perform failure recovery from a single point of control.

GDPS helps automate recovery procedures for planned and unplanned outages to provide near-continuous availability and disaster recovery capability.

For additional information on GDPS, visit:
<http://www-03.ibm.com/systems/z/gdps/>.

Fiber Quick Connect (FQC), an optional feature on z10 EC, is now being offered for all FICON LX (single mode fiber) channels, in addition to the current support for ESCON. FQC is designed to significantly reduce the amount of time required for on-site installation and setup of fiber optic cabling. FQC facilitates adds, moves, and changes of ESCON and FICON LX fiber optic cables in the data center, and may reduce fiber connection time by up to 80%.

FQC is for factory installation of IBM Facilities Cabling Services – Fiber Transport System (FTS) fiber harnesses for connection to channels in the I/O cage. FTS fiber harnesses enable connection to FTS direct-attach fiber trunk cables from IBM Global Technology Services.

Note: FQC supports all of the ESCON channels and all of the FICON LX channels in all of the I/O cages of the server.

z10 EC Physical Characteristics

z10 EC Configuration Detail

z10 EC Environmentals

Model	1 I/O Cage	2 I/O Cage	3 I/O Cage
E12	9.70 kW	13.26 kW	13.50 kW
E26	13.77 kW	17.51 kW	21.17 kW
E40	16.92 kW	20.66 kW	24.40 kW
E56	19.55 kW	23.29 kW	27.00 kW
E64	19.55 kW	23.29 kW	27.50 kW

Model	1 I/O Cage	2 I/O Cage	3 I/O Cage
E12	33.1 kBTU/hr	46.0 kBTU/hr	46.0 kBTU/hr
E26	47.7 kBTU/hr	61.0 kBTU/hr	73.7 kBTU/hr
E40	58.8 kBTU/hr	72.0 kBTU/hr	84.9 kBTU/hr
E56	67.9 kBTU/hr	81.2 kBTU/hr	93.8 kBTU/hr
E64	67.9 kBTU/hr	81.2 kBTU/hr	93.8 kBTU/hr

Note: Model E12 has sufficient Host Channel Adaptor capacity for 58 I/O cards only.

z10 EC Dimensions

	z10 EC	z9 EC
Number of Frames	2 Frame (IBF Contained w/in 2 Frames)	2 Frame (IBF Contained w/in 2 Frames)
Height (with covers)	201.5 cm /79.3 in	194.1 cm /76.4 in
Width (with covers)	156.8 cm /61.7 in	156.8 cm /61.7 in
Depth (with covers)	180.3 cm /71.0 in	157.7 cm /62.1 in
Height Reduction	180.9 cm /71.2 in	178.5 cm /70.3 in
Width Reduction	None	None
Machine Area	2.83 sq. m. /30.44 sq. ft.	2.49 sq. m. /26.78 sq. ft.
Service Clearance (IBF Contained w/in Frame)	5.57 sq. m. /60.00 sq. ft. (IBF Contained w/in Frame)	5.45 sq. m. /58.69 sq. ft. (IBF Contained w/in Frame)

Maximum of 1024 CHPIDs; 3 I/O cages (28 slots each) = 84 I/O slots. All features that require I/O slots, and ICB-4 features, are included in the following table:

Features	Min # Features	Max # Features	Max Connections	Increments per Feature	Purchase Increments
16-port ESCON	0 ⁽¹⁾	69	1024 channels	16 channels 1 reserved as a spare	4 channels
FICON Express4	0 ⁽¹⁾	84	336 channels	4 channels	4 channels
FICON Express2**	0 ⁽¹⁾	84	336 channels	4 channels	4 channels
FICON Express**	0 ⁽¹⁾	60	120 channels	2 channels	2 channels
ICB-4	0 ⁽¹⁾	8	16 links ^{(2) (3)}	2 links	1 link
ISC-3	0 ⁽¹⁾	12	48 links ⁽²⁾	4 links	1 link
HCA2-0 LR (1x)	0 ⁽¹⁾	16	32 links ^{(2) (3)}	2 links	2 links
HCA2-0 (12x)	0 ⁽¹⁾	16	32 links ^{(2) (3)}	2 links	2 links
OSA-Express3*	0	24	48/96 ports	2 or 4	2 ports/ 4 ports
OSA-Express2**	0	24	48 ports	1 or 2	2 ports/ 1 port
Crypto Express2*	0	8	16 PCI-X adapters	2 PCI-X adapters	2 PCI-X adapters ⁽⁴⁾

1. Minimum of one I/O feature (ESCON, FICON) or Coupling Link (PSIFB, ICB-4, ISC-3) required.
 2. The maximum number of external Coupling Links combined cannot exceed 64 per server. There is a maximum of 64 coupling link CHPIDs per server (ICs, ICB-4s, active ISC-3 links, and IFBs)
 3. ICB-4 and 12x IB-DDR are not included in the maximum feature count for I/O slots but are included in the CHPID count.
 4. Initial order of Crypto Express2 is 4 PCI-X adapters (two features). Each PCI-X adapter can be configured as a coprocessor or an accelerator.
- * OSA-Express3 GbE and 1000BASE-T have 2 and 4 port options
- ** Available only when carried forward on an upgrade from z890 or z9 BC. Limited availability for OSA-Express2 GbE features

Processor Unit Features

Model	Books	CPs	IFLs uIFLs	zAAPs zIIPs	ICFs	Standard SAP	Standard Spares
E12	1/17	0-12	0-12 0-11	0-6 0-6	0-12	3	2
E26	2/34	0-26	0-26 0-25	0-13 0-13	0-16	6	2
E40	3/51	0-40	0-40 0-39	0-20 0-20	0-16	9	2
E56	4/68	0-56	0-56 0-55	0-28 0-28	0-16	10	2
E64	4/77	0-64	0-64 0-63	0-32 0-32	0-16	11	2

Note: a minimum of one CP, IFL, or ICF must be purchased on every model.

Note: One zAAP and one zIIP may be purchased for each CP purchased.

Standard memory

z10 EC	Minimum	Maximum
E12	16 GB	352 GB
E26	16 GB	752 GB
E40	16 GB	1136 GB
E56	16 GB	1520 GB
E64	16 GB	1520 GB

Memory cards include: 8 GB, 16 GB, 32 GB, 48 GB and 64 GB. (Fixed HSA not included)

Channels

z10 Model	E12	E26	E40	E56	E64
ESCON Min	0	0	0	0	0
ESCON Max	960	1024	1024	1024	1024
FICON Express4 Min FICON Express2 Min FICON Express Min	0	0	0	0	0
FICON Express4 Max	256	336	336	336	336
FICON Express2 Max*	256	336	336	336	336
FICON Express Max*	120	120	120	120	120

Note: Minimum of one I/O feature (ESCON, FICON) or one Coupling required.

*Available only when carried forward on an upgrade from z9 EC or z990.

Coupling Links

Links	PSIFB	ICB-4	ISC-3	IC	Max Links
	0-32* Except E64	0-16*	0-48	0-32	Total External + Internal Links = 64

* Maximum of 32 IFB + ICB-4 links on System z10 EC. ICB-4 not supported on Model E64.

Cryptographic Features

	Crypto Express2 Feature*
Minimum	0
Maximum	8

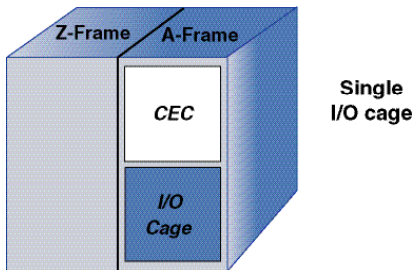
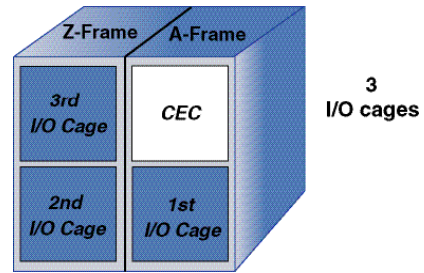
* Each feature has 2 PCI-X adapters; each adapter can be configured as a coprocessor or an accelerator.

OSA-Express3 and OSA-Express2 Features

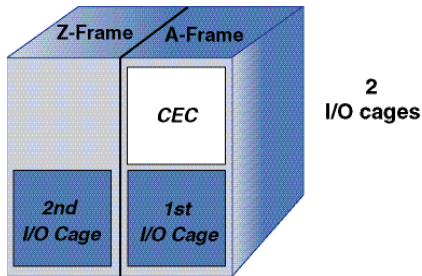
	Min Feat.	Max Feat.	Max Connections	Increments per feat.	Purchase Increments
OSA-Express3	0	24	96	2 ports for 10 GbE	2 ports
OSA-Express2	2	24	48 (10 GbE has 1)	2 or 1	2 ports/ 1 port

z10 EC Frame and I/O Configuration Content: Planning for I/O

The following diagrams show the capability and flexibility built into the I/O subsystem. All machines are shipped with two frames, the A-Frame and the Z-Frame, and can have between one and three I/O cages. Each I/O cage has 28 I/O slots.



I/O Feature Type	Features	Maximum
ESCON	24	360 channels
FICON Express2/4	24	96 channels
FICON Express	24	48 channels
OSA-Express3	24	48/96 (2 or 4 ports)
OSA-Express2	24	48 ports
OSA-Express3 LR/SR	24	48 ports
Crypto Express2	8	16 adapters



I/O Feature Type	Features	Maximum
ESCON	48	720 channels
FICON Express2/4	48	192 channels
FICON Express	48	96 channels
OSA-Express3	24	48/96 (2 or 4 ports)
OSA-Express2	24	48 ports
OSA-Express3 LR/SR	24	48 ports
Crypto Express2	8	16 adapters

I/O Feature Type	Features	Maximum
ESCON	69	1024 channels
FICON Express2/4	84	336 channels
FICON Express	60	120 channels
OSA-Express3	24	48/96 (2 or 4 ports)
OSA-Express2	24	48 ports
OSA-Express3 LR/SR	24	48 ports
Crypto Express2	8	16 adapters

General Information:

- ESCON configured in 4-port increments. Up to a maximum 69 cards, 1024 channels.
- OSA-Express2 can be Gigabit Ethernet (GbE), 1000BASE-T Ethernet or 10 GbE.
- OSA-Express can be Gigabit Ethernet (GbE), 1000BASE-T Ethernet or Fast Ethernet.
- If ICB-3 is required on the system, it will use up a single I/O slot for every 2 ICB-3 to accommodate the STI-3 card.

Note: In the first and second I/O cage, the last domain in the I/O cage is normally used for ISC-3 and ICB-3 links. When the first 6 domains in an I/O cage are full, additional I/O cards will be installed in the next I/O cage. When all the first 6 domains in all I/O cages are full and no Coupling link or PSC cards are required, the last domain in the I/O cage will be used for other I/O cards making a total of 28 per cage.

Coupling Facility – CF Level of Support

CF Level	Function	z10 EC z10 BC	z9 EC z9 BC	z990 z890
16	CF Duplexing Enhancements List Notification Improvements Structure Size increment increase from 512 MB → 1 MB		X	
15	Increasing the allowable tasks in the CF from 48 to 112	X	X	
14	CFCC Dispatcher Enhancements		X	X
13	DB2 Castout Performance		X	X
12	z990 Compatibility 64-bit CFCC Addressability Message Time Ordering DB2 Performance SM Duplexing Support for zSeries		X X X	X X X
11	z990 Compatibility SM Duplexing Support for 9672 G5/G6/R06		X	X
10	z900 GA2 Level		X	X
9	Intelligent Resource Director IC3 / ICB-3 / ISC-3 Peer Mode MQSeries Shared Queues WLM Multi-System Enclaves		X X X	X X X

Note: zSeries 900/800 and prior generation servers are not supported with System z10 for Coupling Facility or Parallel Sysplex levels.

Statement of Direction

IBM intends to support optional water cooling on future high end System z servers. This cooling technology will tap into building chilled water that already exists within the datacenter for computer room air conditioning systems. External chillers or special water conditioning will not be required. Water cooling technology for high end System z servers will be designed to deliver improved energy efficiencies.

IBM intends to support the ability to operate from High Voltage DC power on future System z servers. This will be in addition to the wide range of AC power already supported. A direct HV DC datacenter power design can improve data center energy efficiency by removing the need for an additional DC to AC inversion step.

The System z10 will be the last server to support Dynamic ICF expansion. This is consistent with the System z9 hardware announcement 107-190 dated April 18, 2007, IBM System z9 Enterprise Class (z9 EC) and System z9 Business Class (z9 BC) – Delivering greater value for every-

one, in which the following Statement of Direction was made: IBM intends to remove the Dynamic ICF expansion function from future System z servers.

The System z10 will be the last server to support connections to the Sysplex Timer (9037). Servers that require time synchronization, such as to support a base or Parallel Sysplex, will require Server Time Protocol (STP). STP has been available since January 2007 and is offered on the System z10, System z9, and zSeries 990 and 890 servers.

ESCON channels to be phased out: It is IBM's intent for ESCON channels to be phased out. System z10 EC and System z10 BC will be the last servers to support greater than 240 ESCON channels.

ICB-4 links to be phased out: Restatement of SOD) from RFA46507) IBM intends to not offer Integrated Cluster Bus-4 (ICB-4) links on future servers. IBM intends for System z10 to be the last server to support ICB-4 links.

Publications

The following Redbook publications are available now:

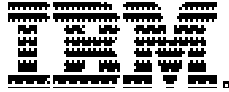
z10 EC Technical Overview SG24-7515		Physical Layer	SA23-0395
z10 EC Technical Guide	SG24-7516	ESCON and FICON CTC Reference	SB10-7034
z10 EC Capacity on Demand	SG24-7504	ESCON I/O Interface Physical Layer	SA23-0394
Getting Started with InfiniBand on z10 EC and System z9	SG24-7539	FICON I/O Interface Physical Layer SA24-7172	
		Hardware Management Console Operations Guide (V2.10.0)	SC28-6867

The following publications are available in the Library section of

Resource Link:

z10 EC System Overview	SA22-1084	IOCP User's Guide	SB10-7037
z10 EC Installation Manual - Physical Planning (IMPP)	GC28-6865	Maintenance Information for Fiber Optic Links	SY27-2597
z10 EC PR/SM Planning Guide	SB10-7153	z10 EC Parts Catalog	GC28-6869
z10 EC Installation Manual	GC28-6864	Planning for Fiber Optic Links	GA23-0367
z10 EC Service Guide	GC28-6866	SCSI IPL - Machine Loader Messages	SC28-6839
z10 EC Safety Inspection Guide	GC28-6870	Service Guide for HMCs and SEs	GC28-6861
System Safety Notices	G229-9054	Service Guide for Trusted Key Entry Workstations	GC28-6862
Application Programming Interfaces for Java	API-JAVA	Standalone IOCP User's Guide	SB10-7152
Application Programming Interfaces Capacity on Demand User's Guide	SB10-7030	Support Element Operations Guide (Version 2.10.0)	SC28-6868
CHPID Mapping Tool User's Guide		System z Functional Matrix	ZSW01335
Common Information Model (CIM) Management Interface	SB10-7154	OSA-Express Customer's Guide	SA22-7935
Coupling Facility Channel I/O Interface		OSA-ICC User's Guide	SA22-7990

Publications for System z10 Enterprise Class can be obtained at Resource Link by accessing the following Web site: www.ibm.com/servers/resourcelink.



© Copyright IBM Corporation 2009

IBM Systems and Technology Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America,
04-09
All Rights Reserved

References in this publication to IBM products or services do not imply that IBM intends to make them available in every country in which IBM operates. Consult your local IBM business contact for information on the products, features, and services available in your area.

IBM, IBM eServer, the IBM logo, the e-business logo, , AIX, APPN, CICS, Cool Blue, DB2, DRDA, DS8000, Dynamic Infrastructure, ECKD, ESCON, FICON, Geographically Dispersed Parallel Sysplex, GDPS, HiperSockets, HyperSwap, IMS, Lotus, MQSeries, MVS, OS/390, Parallel Sysplex, PR/SM, Processor Resource/Systems Manager, RACF, Rational, Redbooks, Resource Link, RETAIN, REXX, RMF, S/390, Scalable Architecture for Financial Reporting, Sysplex Timer, Systems Director Active Energy Manager, System Storage, System z, System z9, System z10, Tivoli, TotalStorage, VSE/ESA, VTAM, WebSphere, z9, z10, z10 BC, z10 EC, z/Architecture, z/OS, z/VM, z/VSE, and zSeries are trademarks or registered trademarks of the International Business Machines Corporation in the United States and other countries.

InfiniBand is a trademark and service mark of the InfiniBand Trade Association.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States or other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel is a trademark of the Intel Corporation in the United States and other countries.

Other trademarks and registered trademarks are the properties of their respective companies.

IBM hardware products are manufactured from new parts, or new and used parts. Regardless, our warranty terms apply.

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

All performance information was determined in a controlled environment. Actual results may vary. Performance information is provided "AS IS" and no warranties or guarantees are expressed or implied by IBM.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

This equipment is subject to all applicable FCC rules and will comply with them upon delivery.

Information concerning non-IBM products was obtained from the suppliers of those products. Questions concerning those products should be directed to those suppliers.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by custom.